

**UTILITY
PATENT APPLICATION
TRANSMITTAL**

Attorney Docket No.

210121.455C13

First Inventor or Application Identifier

Tongtong Wang

Title

**COMPOSITIONS AND METHODS FOR THE THERAPY
AND DIAGNOSIS OF LUNG CANCER**

Express Mail Label No.

EL615232245US

APPLICATION ELEMENTS

See MPEP chapter 600 concerning utility patent application contents.

ADDRESS TO:Box Patent Application
Assistant Commissioner for Patent
Washington, D.C. 202311. ☐ General Authorization Form & Fee Transmittal
(Submit an original and a duplicate for fee processing)2. ☒ Specification [Total Pages] **161**
(preferred arrangement set forth below)

- Descriptive Title of the Invention
- Cross References to Related Applications
- Statement Regarding Fed sponsored R & D
- Reference to Microfiche Appendix
- Background of the Invention
- Brief Summary of the Invention
- Brief Description of the Drawings (if filed)
- Detailed Description
- Claim(s)
- Abstract of the Disclosure

3. ☒ Drawing(s) (35 USC 113) [Total Sheets] **3**4. Oath or Declaration [Total Pages] **1**a. ☐ Newly executed (original or copy)b. ☐ Copy from a prior application (37 CFR 1.63(d))
(for continuation/divisional with Box 17 completed)i. ☐ **DELETION OF INVENTOR(S)**
Signed statement attached deleting
inventor(s) named in the prior application,
see 37 CFR 1.63(d)(2) and 1.33(b)5. ☐ Incorporation By Reference (useable if box 4b is
checked) The entire disclosure of the prior application,
from which a copy of the oath or declaration is supplied
under Box 4b, is considered to be part of the disclosure of
the accompanying application and is hereby incorporated by
reference therein.6. ☐ Microfiche Computer Program (Appendix)7. Nucleotide and Amino Acid Sequence Submission
(if applicable, all necessary)a. ☒ Computer-Readable Copyb. ☒ Paper Copy (identical to computer copy)c. ☒ Statement verifying identity of above copies**ACCOMPANYING APPLICATION PARTS**8. ☐ Assignment Papers (cover sheet & document(s))9. ☐ 37 CFR 3.73(b) Statement ☐ Power of Attorney
(when there is an assignee)10. ☐ English Translation Document (if applicable)11. ☐ Information Disclosure Statement (IDS)/PTO-1449 ☐ Copies of IDS
Citations12. ☐ Preliminary Amendment13. ☒ Return Receipt Postcard14. ☐ Small Entity Statement(s) ☐ Statement filed in prior application,
Status still proper and desired15. ☐ Certified Copy of Priority Document(s)
(if foreign priority is claimed)16. ☒ Other: Certificate of Express Mail

17. If a CONTINUING APPLICATION, check appropriate box and supply the requisite information below and in a preliminary amendment

☐ Continuation ☐ Divisional ☒ Continuation-In-Part (CIP) of prior Application No.: 09/not assignedPrior application information: Examiner not assigned Group / Art Unit not assigned☐ Claims the benefit of Provisional Application No. _____**CORRESPONDENCE ADDRESS**Jane E. R. Potter
Seed Intellectual Property Law Group PLLC
701 Fifth Avenue, Suite 6300
Seattle, Washington 98104-7092
Phone: (206) 622-4900 Fax: (206) 682-6031

Respectfully submitted,

TYPED or PRINTED NAME Jane E. R. PotterSIGNATURE Jane E. R. Potter

u:\sharons\210121\455

REGISTRATION NO. 33,332Date October 9, 2000

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

Applicants : Tongtong Wang, Medina, WA; Liquan Fan, Bellevue, WA; Michael D. Kalos, Seattle, WA; Chaitanya S. Bangur, Seattle, WA; Nancy A. Hosken, Seattle, WA; Gary R. Fanger, Mill Creek, WA; Samuel X. Li; Redmond WA; Aijun Wang, Issaquah, WA; Yasir A. W. Skeiky, Seattle, WA; Robert A. Henderson, Edmonds, WA; Patricia D. McNeill, Des Moines, WA



Filed : October 9, 2000

For : COMPOSITIONS AND METHODS FOR THE THERAPY AND
DIAGNOSIS OF LUNG CANCER

Docket No. : 210121.455C13

Date : October 9, 2000

Box Patent Application
Assistant Commissioner for Patents
Washington, DC 20231

CERTIFICATE OF MAILING BY "EXPRESS MAIL"

Assistant Commissioner for Patents:

I hereby certify that the enclosures listed below are being deposited with the United States Postal Service "EXPRESS MAIL Post Office to Addressee" service under 37 C.F.R. § 1.10, Mailing Label Certificate No. EL615232245US, on October 9, 2000, addressed to Box Patent Application, Assistant Commissioner for Patents, Washington, DC 20231.

Respectfully submitted,

Seed Intellectual Property Law Group PLLC

Steve Plante/Jeanette West/Susan Johnson

JEP:sds

Enclosures:

- Postcard
- Form PTO/SB/05
- Specification, Claims, Abstract (161 pages)
- 3 Sheets Drawings (Figs. 1-3)
- Sequence Listing (187 pages)
- Declaration for Sequence Listing
- Diskette for Sequence Listing

COMPOSITIONS AND METHODS FOR THE THERAPY AND DIAGNOSIS OF LUNG CANCER

CROSS-REFERENCE TO RELATED APPLICATIONS

5 The present application is related to U.S. Patent Application Nos. _____, filed September 26, 2000; 09/643,597, filed August 21, 2000; 09/630,940 filed August 2, 2000; 09/606,421 filed June 28, 2000; 09/542,615 filed April 4, 2000; 09/510,376 filed February 22, 2000; 09/480,884 filed January 10, 2000; 09/476,496 filed December 30, 1999; 09/466,396 filed December 17, 1999; 09/285,479 filed April 2, 1999; 10 09/221,107 filed December 22, 1998; 09/123,912 filed July 27, 1998; 09/040,802 filed March 18, 1998; each a CIP of the previous application and each pending; and PCT Nos. US99/05798 filed March 17, 1999, published, and US00/08896 filed April 4, 2000, pending; all incorporated by reference herein.

TECHNICAL FIELD OF THE INVENTION

15 The present invention relates generally to therapy and diagnosis of cancer, such as lung cancer. The invention is more specifically related to polypeptides comprising at least a portion of a lung tumor protein, and to polynucleotides encoding such polypeptides. Such polypeptides and polynucleotides may be used in vaccines and pharmaceutical compositions for prevention and treatment of lung cancer, and for the 20 diagnosis and monitoring of such cancers.

BACKGROUND OF THE INVENTION

 Lung cancer is the primary cause of cancer death among both men and women in the U.S., with an estimated 172,000 new cases being reported in 1994. The five-year survival rate among all lung cancer patients, regardless of the stage of disease at 25 diagnosis, is only 13%. This contrasts with a five-year survival rate of 46% among cases detected while the disease is still localized. However, only 16% of lung cancers are discovered before the disease has spread.

Early detection is difficult since clinical symptoms are often not seen until the disease has reached an advanced stage. Currently, diagnosis is aided by the use of chest x-rays, analysis of the type of cells contained in sputum and fiberoptic examination of the bronchial passages. Treatment regimens are determined by the type and stage of the cancer, and include surgery, radiation therapy and/or chemotherapy. In spite of considerable research into therapies for the disease, lung cancer remains difficult to treat.

Accordingly, there remains a need in the art for improved vaccines, treatment methods and diagnostic techniques for lung cancer.

SUMMARY OF THE INVENTION

Briefly stated, the present invention provides compositions and methods for the diagnosis and therapy of cancer, such as lung cancer. In one aspect, the present invention provides polypeptides comprising at least a portion of a lung tumor protein, or a variant thereof. Certain portions and other variants are immunogenic, such that the ability of the variant to react with antigen-specific antisera is not substantially diminished. Within certain embodiments, the polypeptide comprises a sequence that is encoded by a polynucleotide sequence selected from the group consisting of: (a) sequences recited in any one of SEQ ID NO: 1-3, 6-8, 10-13, 15-27, 29, 30, 32, 34-49, 51, 52, 54, 55, 57-59, 61-69, 71, 73, 74, 77, 78, 80-82, 84, 86-96, 107-109, 111, 113, 125, 127, 128, 129, 131-133, 142, 144, 148-151, 153, 154, 157, 158, 160, 167, 168, 171, 179, 182, 184-186, 188-191, 193, 194, 198-207, 209, 210, 213, 214, 217, 220-224, 253-337, 345, 347 and 349; (b) variants of a sequence recited in any one of SEQ ID NO: 1-3, 6-8, 10-13, 15-27, 29, 30, 32, 34-49, 51, 52, 54, 55, 57-59, 61-69, 71, 73, 74, 77, 78, 80-82, 84, 86-96, 107-109, 111, 113, 125, 127, 128, 129, 131-133, 142, 144, 148-151, 153, 154, 157, 158, 160, 167, 168, 171, 179, 182, 184-186, 188-191, 193, 194, 198-207, 209, 210, 213, 214, 217, 220-224, 253-337, 345, 347 and 349; and (c) complements of a sequence of (a) or (b). In specific embodiments, the polypeptides of the present invention comprise at least a portion of a tumor protein that includes an amino acid sequence selected from the group consisting of sequences recited in

any one of SEQ ID NO: 152, 155, 156, 165, 166, 169, 170, 172, 174, 176, 226-252, 338-344 and 346, and variants thereof.

The present invention further provides polynucleotides that encode a polypeptide as described above, or a portion thereof (such as a portion encoding at least 15 amino acid residues of a lung tumor protein), expression vectors comprising such polynucleotides and host cells transformed or transfected with such expression vectors.

Within other aspects, the present invention provides pharmaceutical compositions comprising a polypeptide or polynucleotide as described above and a physiologically acceptable carrier.

Within a related aspect of the present invention, vaccines for prophylactic or therapeutic use are provided. Such vaccines comprise a polypeptide or polynucleotide as described above and an immunostimulant.

The present invention further provides pharmaceutical compositions that comprise: (a) an antibody or antigen-binding fragment thereof that specifically binds to a lung tumor protein; and (b) a physiologically acceptable carrier.

Within further aspects, the present invention provides pharmaceutical compositions comprising: (a) an antigen presenting cell that expresses a polypeptide as described above and (b) a pharmaceutically acceptable carrier or excipient. Antigen presenting cells include dendritic cells, macrophages, monocytes, fibroblasts and B cells.

Within related aspects, vaccines are provided that comprise: (a) an antigen presenting cell that expresses a polypeptide as described above, and (b) an immunostimulant.

The present invention further provides, in other aspects, fusion proteins that comprise at least one polypeptide as described above, as well as polynucleotides encoding such fusion proteins.

Within related aspects, pharmaceutical compositions comprising a fusion protein, or a polynucleotide encoding a fusion protein, in combination with a physiologically acceptable carrier are provided.

Vaccines are further provided, within other aspects, that comprise a fusion protein, or a polynucleotide encoding a fusion protein, in combination with an immunostimulant.

Within further aspects, the present invention provides methods for inhibiting the development of a cancer in a patient, comprising administering to a patient a pharmaceutical composition or vaccine as recited above.

The present invention further provides, within other aspects, methods for removing tumor cells from a biological sample, comprising contacting a biological sample with T cells that specifically react with a lung tumor protein, wherein the step of contacting is performed under conditions and for a time sufficient to permit the removal of cells expressing the protein from the sample.

Within related aspects, methods are provided for inhibiting the development of a cancer in a patient, comprising administering to a patient a biological sample treated as described above.

Methods are further provided, within other aspects, for stimulating and/or expanding T cells specific for a lung tumor protein, comprising contacting T cells with one or more of: (i) a polypeptide as described above; (ii) a polynucleotide encoding such a polypeptide; and/or (iii) an antigen presenting cell that expresses such a polypeptide; under conditions and for a time sufficient to permit the stimulation and/or expansion of T cells. Determined T cell populations comprising T cells prepared as described above are also provided.

Within further aspects, the present invention provides methods for inhibiting the development of a cancer in a patient, comprising administering to a patient an effective amount of a T cell population as described above.

The present invention further provides methods for inhibiting the development of a cancer in a patient, comprising the steps of: (a) incubating CD4⁺ and/or CD8⁺ T cells determined from a patient with one or more of: (i) a polypeptide comprising at least an immunogenic portion of a lung tumor protein; (ii) a polynucleotide encoding such a polypeptide; and (iii) an antigen-presenting cell that expressed such a polypeptide;

and (b) administering to the patient an effective amount of the proliferated T cells, and thereby inhibiting the development of a cancer in the patient. Proliferated cells may, but need not, be cloned prior to administration to the patient.

Within further aspects, the present invention provides methods for
 5 determining the presence or absence of a cancer in a patient, comprising: (a) contacting a biological sample obtained from a patient with a binding agent that binds to a polypeptide as recited above; (b) detecting in the sample an amount of polypeptide that binds to the binding agent; and (c) comparing the amount of polypeptide with a predetermined cut-off value, and therefrom determining the presence or absence of a cancer in the patient.
 10 Within preferred embodiments, the binding agent is an antibody, more preferably a monoclonal antibody. The cancer may be lung cancer.

The present invention also provides, within other aspects, methods for monitoring the progression of a cancer in a patient. Such methods comprise the steps of:
 15 (a) contacting a biological sample obtained from a patient at a first point in time with a binding agent that binds to a polypeptide as recited above; (b) detecting in the sample an amount of polypeptide that binds to the binding agent; (c) repeating steps (a) and (b) using a biological sample obtained from the patient at a subsequent point in time; and (d) comparing the amount of polypeptide detected in step (c) with the amount detected in step (b) and therefrom monitoring the progression of the cancer in the patient.

20 The present invention further provides, within other aspects, methods for determining the presence or absence of a cancer in a patient, comprising the steps of: (a) contacting a biological sample obtained from a patient with an oligonucleotide that hybridizes to a polynucleotide that encodes a lung tumor protein; (b) detecting in the sample a level of a polynucleotide, preferably mRNA, that hybridizes to the
 25 oligonucleotide; and (c) comparing the level of polynucleotide that hybridizes to the oligonucleotide with a predetermined cut-off value, and therefrom determining the presence or absence of a cancer in the patient. Within certain embodiments, the amount of mRNA is detected via polymerase chain reaction using, for example, at least one oligonucleotide primer that hybridizes to a polynucleotide encoding a polypeptide as

recited above, or a complement of such a polynucleotide. Within other embodiments, the amount of mRNA is detected using a hybridization technique, employing an oligonucleotide probe that hybridizes to a polynucleotide that encodes a polypeptide as recited above, or a complement of such a polynucleotide.

5 In related aspects, methods are provided for monitoring the progression of a cancer in a patient, comprising the steps of: (a) contacting a biological sample obtained from a patient with an oligonucleotide that hybridizes to a polynucleotide that encodes a lung tumor protein; (b) detecting in the sample an amount of a polynucleotide that hybridizes to the oligonucleotide; (c) repeating steps (a) and (b) using a biological sample
10 obtained from the patient at a subsequent point in time; and (d) comparing the amount of polynucleotide detected in step (c) with the amount detected in step (b) and therefrom monitoring the progression of the cancer in the patient.

Within further aspects, the present invention provides antibodies, such as monoclonal antibodies, that bind to a polypeptide as described above, as well as diagnostic
15 kits comprising such antibodies. Diagnostic kits comprising one or more oligonucleotide probes or primers as described above are also provided.

These and other aspects of the present invention will become apparent upon reference to the following detailed description. All references disclosed herein are hereby incorporated by reference in their entirety as if each was incorporated individually.

20

BRIEF DESCRIPTION OF THE FIGURES AND SEQUENCE IDENTIFIERS

SEQ ID NO: 1 is the determined cDNA sequence for LST-S1-2

SEQ ID NO: 2 is the determined cDNA sequence for LST-S1-28

SEQ ID NO: 3 is the determined cDNA sequence for LST-S1-90

25 SEQ ID NO: 4 is the determined cDNA sequence for LST-S1-144

SEQ ID NO: 5 is the determined cDNA sequence for LST-S1-133

SEQ ID NO: 6 is the determined cDNA sequence for LST-S1-169

SEQ ID NO: 7 is the determined cDNA sequence for LST-S2-6

- SEQ ID NO: 8 is the determined cDNA sequence for LST-S2-11
- SEQ ID NO: 9 is the determined cDNA sequence for LST-S2-17
- SEQ ID NO: 10 is the determined cDNA sequence for LST-S2-25
- SEQ ID NO: 11 is the determined cDNA sequence for LST-S2-39
- 5 SEQ ID NO: 12 is a first determined cDNA sequence for LST-S2-43
- SEQ ID NO: 13 is a second determined cDNA sequence for LST-S2-43
- SEQ ID NO: 14 is the determined cDNA sequence for LST-S2-65
- SEQ ID NO: 15 is the determined cDNA sequence for LST-S2-68
- SEQ ID NO: 16 is the determined cDNA sequence for LST-S2-72
- 10 SEQ ID NO: 17 is the determined cDNA sequence for LST-S2-74
- SEQ ID NO: 18 is the determined cDNA sequence for LST-S2-103
- SEQ ID NO: 19 is the determined cDNA sequence for LST-S2-N1-1F
- SEQ ID NO: 20 is the determined cDNA sequence for LST-S2-N1-2A
- SEQ ID NO: 21 is the determined cDNA sequence for LST-S2-N1-4H
- 15 SEQ ID NO: 22 is the determined cDNA sequence for LST-S2-N1-5A
- SEQ ID NO: 23 is the determined cDNA sequence for LST-S2-N1-6B
- SEQ ID NO: 24 is the determined cDNA sequence for LST-S2-N1-7B
- SEQ ID NO: 25 is the determined cDNA sequence for LST-S2-N1-7H
- SEQ ID NO: 26 is the determined cDNA sequence for LST-S2-N1-8A
- 20 SEQ ID NO: 27 is the determined cDNA sequence for LST-S2-N1-8D
- SEQ ID NO: 28 is the determined cDNA sequence for LST-S2-N1-9A
- SEQ ID NO: 29 is the determined cDNA sequence for LST-S2-N1-9E
- SEQ ID NO: 30 is the determined cDNA sequence for LST-S2-N1-10A
- SEQ ID NO: 31 is the determined cDNA sequence for LST-S2-N1-10G
- 25 SEQ ID NO: 32 is the determined cDNA sequence for LST-S2-N1-11A
- SEQ ID NO: 33 is the determined cDNA sequence for LST-S2-N1-12C
- SEQ ID NO: 34 is the determined cDNA sequence for LST-S2-N1-12E
- SEQ ID NO: 35 is the determined cDNA sequence for LST-S2-B1-3D
- SEQ ID NO: 36 is the determined cDNA sequence for LST-S2-B1-6C

- SEQ ID NO: 37 is the determined cDNA sequence for LST-S2-B1-5D
 SEQ ID NO: 38 is the determined cDNA sequence for LST-S2-B1-5F
 SEQ ID NO: 39 is the determined cDNA sequence for LST-S2-B1-6G
 SEQ ID NO: 40 is the determined cDNA sequence for LST-S2-B1-8A
 5 SEQ ID NO: 41 is the determined cDNA sequence for LST-S2-B1-8D
 SEQ ID NO: 42 is the determined cDNA sequence for LST-S2-B1-10A
 SEQ ID NO: 43 is the determined cDNA sequence for LST-S2-B1-9B
 SEQ ID NO: 44 is the determined cDNA sequence for LST-S2-B1-9F
 SEQ ID NO: 45 is the determined cDNA sequence for LST-S2-B1-12D
 10 SEQ ID NO: 46 is the determined cDNA sequence for LST-S2-I2-2B
 SEQ ID NO: 47 is the determined cDNA sequence for LST-S2-I2-5F
 SEQ ID NO: 48 is the determined cDNA sequence for LST-S2-I2-6B
 SEQ ID NO: 49 is the determined cDNA sequence for LST-S2-I2-7F
 SEQ ID NO: 50 is the determined cDNA sequence for LST-S2-I2-8G
 15 SEQ ID NO: 51 is the determined cDNA sequence for LST-S2-I2-9E
 SEQ ID NO: 52 is the determined cDNA sequence for LST-S2-I2-12B
 SEQ ID NO: 53 is the determined cDNA sequence for LST-S2-H2-2C
 SEQ ID NO: 54 is the determined cDNA sequence for LST-S2-H2-1G
 SEQ ID NO: 55 is the determined cDNA sequence for LST-S2-H2-4G
 20 SEQ ID NO: 56 is the determined cDNA sequence for LST-S2-H2-3H
 SEQ ID NO: 57 is the determined cDNA sequence for LST-S2-H2-5G
 SEQ ID NO: 58 is the determined cDNA sequence for LST-S2-H2-9B
 SEQ ID NO: 59 is the determined cDNA sequence for LST-S2-H2-10H
 SEQ ID NO: 60 is the determined cDNA sequence for LST-S2-H2-12D
 25 SEQ ID NO: 61 is the determined cDNA sequence for LST-S3-2
 SEQ ID NO: 62 is the determined cDNA sequence for LST-S3-4
 SEQ ID NO: 63 is the determined cDNA sequence for LST-S3-7
 SEQ ID NO: 64 is the determined cDNA sequence for LST-S3-8
 SEQ ID NO: 65 is the determined cDNA sequence for LST-S3-12

- SEQ ID NO: 66 is the determined cDNA sequence for LST-S3-13
 SEQ ID NO: 67 is the determined cDNA sequence for LST-S3-14
 SEQ ID NO: 68 is the determined cDNA sequence for LST-S3-16
 SEQ ID NO: 69 is the determined cDNA sequence for LST-S3-21
 5 SEQ ID NO: 70 is the determined cDNA sequence for LST-S3-22
 SEQ ID NO: 71 is the determined cDNA sequence for LST-S1-7
 SEQ ID NO: 72 is the determined cDNA sequence for LST-S1-A-1E
 SEQ ID NO: 73 is the determined cDNA sequence for LST-S1-A-1G
 SEQ ID NO: 74 is the determined cDNA sequence for LST-S1-A-3E
 10 SEQ ID NO: 75 is the determined cDNA sequence for LST-S1-A-4E
 SEQ ID NO: 76 is the determined cDNA sequence for LST-S1-A-6D
 SEQ ID NO: 77 is the determined cDNA sequence for LST-S1-A-8D
 SEQ ID NO: 78 is the determined cDNA sequence for LST-S1-A-10A
 SEQ ID NO: 79 is the determined cDNA sequence for LST-S1-A-10C
 15 SEQ ID NO: 80 is the determined cDNA sequence for LST-S1-A-9D
 SEQ ID NO: 81 is the determined cDNA sequence for LST-S1-A-10D
 SEQ ID NO: 82 is the determined cDNA sequence for LST-S1-A-9H
 SEQ ID NO: 83 is the determined cDNA sequence for LST-S1-A-11D
 SEQ ID NO: 84 is the determined cDNA sequence for LST-S1-A-12D
 20 SEQ ID NO: 85 is the determined cDNA sequence for LST-S1-A-11E
 SEQ ID NO: 86 is the determined cDNA sequence for LST-S1-A-12E
 SEQ ID NO: 87 is the determined cDNA sequence for L513S (T3).
 SEQ ID NO: 88 is the determined cDNA sequence for L513S contig 1.
 SEQ ID NO: 89 is a first determined cDNA sequence for L514S.
 25 SEQ ID NO: 90 is a second determined cDNA sequence for L514S.
 SEQ ID NO: 91 is a first determined cDNA sequence for L516S.
 SEQ ID NO: 92 is a second determined cDNA sequence for L516S.
 SEQ ID NO: 93 is the determined cDNA sequence for L517S.
 SEQ ID NO: 94 is the extended cDNA sequence for LST-S1-169 (also known as L519S).

- SEQ ID NO: 95 is a first determined cDNA sequence for L520S.
- SEQ ID NO: 96 is a second determined cDNA sequence for L520S.
- SEQ ID NO: 97 is a first determined cDNA sequence for L521S.
- SEQ ID NO: 98 is a second determined cDNA sequence for L521S.
- 5 SEQ ID NO: 99 is the determined cDNA sequence for L522S.
- SEQ ID NO: 100 is the determined cDNA sequence for L523S.
- SEQ ID NO: 101 is the determined cDNA sequence for L524S.
- SEQ ID NO: 102 is the determined cDNA sequence for L525S.
- SEQ ID NO: 103 is the determined cDNA sequence for L526S.
- 10 SEQ ID NO: 104 is the determined cDNA sequence for L527S.
- SEQ ID NO: 105 is the determined cDNA sequence for L528S.
- SEQ ID NO: 106 is the determined cDNA sequence for L529S.
- SEQ ID NO: 107 is a first determined cDNA sequence for L530S.
- SEQ ID NO: 108 is a second determined cDNA sequence for L530S.
- 15 SEQ ID NO: 109 is the determined full-length cDNA sequence for L531S short form
- SEQ ID NO: 110 is the predicted amino acid sequence encoded by SEQ ID NO: 109.
- SEQ ID NO: 111 is the determined full-length cDNA sequence for L531S long form
- SEQ ID NO: 112 is the predicted amino acid sequence encoded by SEQ ID NO: 111.
- SEQ ID NO: 113 is the determined full-length cDNA sequence for L520S.
- 20 SEQ ID NO: 114 is the predicted amino acid sequence encoded by SEQ ID NO: 113.
- SEQ ID NO: 115 is the determined cDNA sequence for contig 1.
- SEQ ID NO: 116 is the determined cDNA sequence for contig 3.
- SEQ ID NO: 117 is the determined cDNA sequence for contig 4.
- SEQ ID NO: 118 is the determined cDNA sequence for contig 5.
- 25 SEQ ID NO: 119 is the determined cDNA sequence for contig 7.
- SEQ ID NO: 120 is the determined cDNA sequence for contig 8.
- SEQ ID NO: 121 is the determined cDNA sequence for contig 9.
- SEQ ID NO: 122 is the determined cDNA sequence for contig 10.
- SEQ ID NO: 123 is the determined cDNA sequence for contig 12.

SEQ ID NO: 124 is the determined cDNA sequence for contig 11.

SEQ ID NO: 125 is the determined cDNA sequence for contig 13 (also known as L761P).

SEQ ID NO: 126 is the determined cDNA sequence for contig 15.

SEQ ID NO: 127 is the determined cDNA sequence for contig 16.

5 SEQ ID NO: 128 is the determined cDNA sequence for contig 17.

SEQ ID NO: 129 is the determined cDNA sequence for contig 19.

SEQ ID NO: 130 is the determined cDNA sequence for contig 20.

SEQ ID NO: 131 is the determined cDNA sequence for contig 22.

SEQ ID NO: 132 is the determined cDNA sequence for contig 24.

10 SEQ ID NO: 133 is the determined cDNA sequence for contig 29.

SEQ ID NO: 134 is the determined cDNA sequence for contig 31.

SEQ ID NO: 135 is the determined cDNA sequence for contig 33.

SEQ ID NO: 136 is the determined cDNA sequence for contig 38.

SEQ ID NO: 137 is the determined cDNA sequence for contig 39.

15 SEQ ID NO: 138 is the determined cDNA sequence for contig 41.

SEQ ID NO: 139 is the determined cDNA sequence for contig 43.

SEQ ID NO: 140 is the determined cDNA sequence for contig 44.

SEQ ID NO: 141 is the determined cDNA sequence for contig 45.

SEQ ID NO: 142 is the determined cDNA sequence for contig 47.

20 SEQ ID NO: 143 is the determined cDNA sequence for contig 48.

SEQ ID NO: 144 is the determined cDNA sequence for contig 49.

SEQ ID NO: 145 is the determined cDNA sequence for contig 50.

SEQ ID NO: 146 is the determined cDNA sequence for contig 53.

SEQ ID NO: 147 is the determined cDNA sequence for contig 54.

25 SEQ ID NO: 148 is the determined cDNA sequence for contig 56.

SEQ ID NO: 149 is the determined cDNA sequence for contig 57.

SEQ ID NO: 150 is the determined cDNA sequence for contig 58.

SEQ ID NO: 151 is the full-length cDNA sequence for L530S.

SEQ ID NO: 152 is the amino acid sequence encoded by SEQ ID NO: 151

SEQ ID NO: 153 is the full-length cDNA sequence of a first variant of L514S

SEQ ID NO: 154 is the full-length cDNA sequence of a second variant of L514S

SEQ ID NO: 155 is the amino acid sequence encoded by SEQ ID NO: 153.

SEQ ID NO: 156 is the amino acid sequence encoded by SEQ ID NO: 154.

5 SEQ ID NO: 157 is the determined cDNA sequence for contig 59.

SEQ ID NO: 158 is the full-length cDNA sequence for L763P (also referred to as contig 22).

SEQ ID NO: 159 is the amino acid sequence encoded by SEQ ID NO: 158.

10 SEQ ID NO: 160 is the full-length cDNA sequence for L762P (also referred to as contig 17).

SEQ ID NO: 161 is the amino acid sequence encoded by SEQ ID NO: 160.

SEQ ID NO: 162 is the determined cDNA sequence for L515S.

SEQ ID NO: 163 is the full-length cDNA sequence of a first variant of L524S.

SEQ ID NO: 164 is the full-length cDNA sequence of a second variant of L524S.

15 SEQ ID NO: 165 is the amino acid sequence encoded by SEQ ID NO: 163.

SEQ ID NO: 166 is the amino acid sequence encoded by SEQ ID NO: 164.

SEQ ID NO: 167 is the full-length cDNA sequence of a first variant of L762P.

SEQ ID NO: 168 is the full-length cDNA sequence of a second variant of L762P.

SEQ ID NO: 169 is the amino acid sequence encoded by SEQ ID NO: 167.

20 SEQ ID NO: 170 is the amino acid sequence encoded by SEQ ID NO: 168.

SEQ ID NO: 171 is the full-length cDNA sequence for L773P (also referred to as contig 56).

SEQ ID NO: 172 is the amino acid sequence encoded by SEQ ID NO: 171.

SEQ ID NO: 173 is an extended cDNA sequence for L519S.

25 SEQ ID NO: 174 is the predicted amino acid sequence encoded by SEQ ID NO: 174.

SEQ ID NO: 175 is the full-length cDNA sequence for L523S.

SEQ ID NO: 176 is the predicted amino acid sequence encoded by SEQ ID NO: 175.

SEQ ID NO: 177 is the determined cDNA sequence for LST-sub5-7A.

SEQ ID NO: 178 is the determined cDNA sequence for LST-sub5-8G.

- SEQ ID NO: 179 is the determined cDNA sequence for LST-sub5-8H.
- SEQ ID NO: 180 is the determined cDNA sequence for LST-sub5-10B.
- SEQ ID NO: 181 is the determined cDNA sequence for LST-sub5-10H.
- SEQ ID NO: 182 is the determined cDNA sequence for LST-sub5-12B.
- 5 SEQ ID NO: 183 is the determined cDNA sequence for LST-sub5-11C.
- SEQ ID NO: 184 is the determined cDNA sequence for LST-sub6-1c.
- SEQ ID NO: 185 is the determined cDNA sequence for LST-sub6-2f.
- SEQ ID NO: 186 is the determined cDNA sequence for LST-sub6-2G.
- SEQ ID NO: 187 is the determined cDNA sequence for LST-sub6-4d.
- 10 SEQ ID NO: 188 is the determined cDNA sequence for LST-sub6-4e.
- SEQ ID NO: 189 is the determined cDNA sequence for LST-sub6-4f.
- SEQ ID NO: 190 is the determined cDNA sequence for LST-sub6-3h.
- SEQ ID NO: 191 is the determined cDNA sequence for LST-sub6-5d.
- SEQ ID NO: 192 is the determined cDNA sequence for LST-sub6-5h.
- 15 SEQ ID NO: 193 is the determined cDNA sequence for LST-sub6-6h.
- SEQ ID NO: 194 is the determined cDNA sequence for LST-sub6-7a.
- SEQ ID NO: 195 is the determined cDNA sequence for LST-sub6-8a.
- SEQ ID NO: 196 is the determined cDNA sequence for LST-sub6-7d.
- SEQ ID NO: 197 is the determined cDNA sequence for LST-sub6-7e.
- 20 SEQ ID NO: 198 is the determined cDNA sequence for LST-sub6-8e.
- SEQ ID NO: 199 is the determined cDNA sequence for LST-sub6-7g.
- SEQ ID NO: 200 is the determined cDNA sequence for LST-sub6-9f.
- SEQ ID NO: 201 is the determined cDNA sequence for LST-sub6-9h.
- SEQ ID NO: 202 is the determined cDNA sequence for LST-sub6-11b.
- 25 SEQ ID NO: 203 is the determined cDNA sequence for LST-sub6-11c.
- SEQ ID NO: 204 is the determined cDNA sequence for LST-sub6-12c.
- SEQ ID NO: 205 is the determined cDNA sequence for LST-sub6-12e.
- SEQ ID NO: 206 is the determined cDNA sequence for LST-sub6-12f.
- SEQ ID NO: 207 is the determined cDNA sequence for LST-sub6-11g.

SEQ ID NO: 208 is the determined cDNA sequence for LST-sub6-12g.

SEQ ID NO: 209 is the determined cDNA sequence for LST-sub6-12h.

SEQ ID NO: 210 is the determined cDNA sequence for LST-sub6-II-1a.

SEQ ID NO: 211 is the determined cDNA sequence for LST-sub6-II-2b.

5 SEQ ID NO: 212 is the determined cDNA sequence for LST-sub6-II-2g.

SEQ ID NO: 213 is the determined cDNA sequence for LST-sub6-II-1h.

SEQ ID NO: 214 is the determined cDNA sequence for LST-sub6-II-4a.

SEQ ID NO: 215 is the determined cDNA sequence for LST-sub6-II-4b.

SEQ ID NO: 216 is the determined cDNA sequence for LST-sub6-II-3e.

10 SEQ ID NO: 217 is the determined cDNA sequence for LST-sub6-II-4f.

SEQ ID NO: 218 is the determined cDNA sequence for LST-sub6-II-4g.

SEQ ID NO: 219 is the determined cDNA sequence for LST-sub6-II-4h.

SEQ ID NO: 220 is the determined cDNA sequence for LST-sub6-II-5c.

SEQ ID NO: 221 is the determined cDNA sequence for LST-sub6-II-5e.

15 SEQ ID NO: 222 is the determined cDNA sequence for LST-sub6-II-6f.

SEQ ID NO: 223 is the determined cDNA sequence for LST-sub6-II-5g.

SEQ ID NO: 224 is the determined cDNA sequence for LST-sub6-II-6g.

SEQ ID NO: 225 is the amino acid sequence for L528S.

SEQ ID NO: 226-251 are synthetic peptides derived from L762P.

20 SEQ ID NO: 252 is the expressed amino acid sequence of L514S.

SEQ ID NO: 253 is the DNA sequence corresponding to SEQ ID NO: 252.

SEQ ID NO: 254 is the DNA sequence of a L762P expression construct.

SEQ ID NO: 255 is the determined cDNA sequence for clone 23785.

SEQ ID NO: 256 is the determined cDNA sequence for clone 23786.

25 SEQ ID NO: 257 is the determined cDNA sequence for clone 23788.

SEQ ID NO: 258 is the determined cDNA sequence for clone 23790.

SEQ ID NO: 259 is the determined cDNA sequence for clone 23793.

SEQ ID NO: 260 is the determined cDNA sequence for clone 23794.

SEQ ID NO: 261 is the determined cDNA sequence for clone 23795.

- SEQ ID NO: 262 is the determined cDNA sequence for clone 23796.
 SEQ ID NO: 263 is the determined cDNA sequence for clone 23797.
 SEQ ID NO: 264 is the determined cDNA sequence for clone 23798.
 SEQ ID NO: 265 is the determined cDNA sequence for clone 23799.
- 5 SEQ ID NO: 266 is the determined cDNA sequence for clone 23800.
 SEQ ID NO: 267 is the determined cDNA sequence for clone 23802.
 SEQ ID NO: 268 is the determined cDNA sequence for clone 23803.
 SEQ ID NO: 269 is the determined cDNA sequence for clone 23804.
 SEQ ID NO: 270 is the determined cDNA sequence for clone 23805.
- 10 SEQ ID NO: 271 is the determined cDNA sequence for clone 23806.
 SEQ ID NO: 272 is the determined cDNA sequence for clone 23807.
 SEQ ID NO: 273 is the determined cDNA sequence for clone 23808.
 SEQ ID NO: 274 is the determined cDNA sequence for clone 23809.
 SEQ ID NO: 275 is the determined cDNA sequence for clone 23810.
- 15 SEQ ID NO: 276 is the determined cDNA sequence for clone 23811.
 SEQ ID NO: 277 is the determined cDNA sequence for clone 23812.
 SEQ ID NO: 278 is the determined cDNA sequence for clone 23813.
 SEQ ID NO: 279 is the determined cDNA sequence for clone 23815.
 SEQ ID NO: 280 is the determined cDNA sequence for clone 25298.
- 20 SEQ ID NO: 281 is the determined cDNA sequence for clone 25299.
 SEQ ID NO: 282 is the determined cDNA sequence for clone 25300.
 SEQ ID NO: 283 is the determined cDNA sequence for clone 25301
 SEQ ID NO: 284 is the determined cDNA sequence for clone 25304
 SEQ ID NO: 285 is the determined cDNA sequence for clone 25309.
- 25 SEQ ID NO: 286 is the determined cDNA sequence for clone 25312.
 SEQ ID NO: 287 is the determined cDNA sequence for clone 25317.
 SEQ ID NO: 288 is the determined cDNA sequence for clone 25321.
 SEQ ID NO: 289 is the determined cDNA sequence for clone 25323.
 SEQ ID NO: 290 is the determined cDNA sequence for clone 25327.

- SEQ ID NO: 291 is the determined cDNA sequence for clone 25328.
- SEQ ID NO: 292 is the determined cDNA sequence for clone 25332.
- SEQ ID NO: 293 is the determined cDNA sequence for clone 25333.
- SEQ ID NO: 294 is the determined cDNA sequence for clone 25336.
- 5 SEQ ID NO: 295 is the determined cDNA sequence for clone 25340.
- SEQ ID NO: 296 is the determined cDNA sequence for clone 25342.
- SEQ ID NO: 297 is the determined cDNA sequence for clone 25356.
- SEQ ID NO: 298 is the determined cDNA sequence for clone 25357.
- SEQ ID NO: 299 is the determined cDNA sequence for clone 25361.
- 10 SEQ ID NO: 300 is the determined cDNA sequence for clone 25363.
- SEQ ID NO: 301 is the determined cDNA sequence for clone 25397.
- SEQ ID NO: 302 is the determined cDNA sequence for clone 25402.
- SEQ ID NO: 303 is the determined cDNA sequence for clone 25403.
- SEQ ID NO: 304 is the determined cDNA sequence for clone 25405.
- 15 SEQ ID NO: 305 is the determined cDNA sequence for clone 25407.
- SEQ ID NO: 306 is the determined cDNA sequence for clone 25409.
- SEQ ID NO: 307 is the determined cDNA sequence for clone 25396.
- SEQ ID NO: 308 is the determined cDNA sequence for clone 25414.
- SEQ ID NO: 309 is the determined cDNA sequence for clone 25410.
- 20 SEQ ID NO: 310 is the determined cDNA sequence for clone 25406.
- SEQ ID NO: 311 is the determined cDNA sequence for clone 25306.
- SEQ ID NO: 312 is the determined cDNA sequence for clone 25362.
- SEQ ID NO: 313 is the determined cDNA sequence for clone 25360.
- SEQ ID NO: 314 is the determined cDNA sequence for clone 25398.
- 25 SEQ ID NO: 315 is the determined cDNA sequence for clone 25355.
- SEQ ID NO: 316 is the determined cDNA sequence for clone 25351.
- SEQ ID NO: 317 is the determined cDNA sequence for clone 25331.
- SEQ ID NO: 318 is the determined cDNA sequence for clone 25338.
- SEQ ID NO: 319 is the determined cDNA sequence for clone 25335.

- SEQ ID NO: 320 is the determined cDNA sequence for clone 25329.
- SEQ ID NO: 321 is the determined cDNA sequence for clone 25324.
- SEQ ID NO: 322 is the determined cDNA sequence for clone 25322.
- SEQ ID NO: 323 is the determined cDNA sequence for clone 25319.
- 5 SEQ ID NO: 324 is the determined cDNA sequence for clone 25316.
- SEQ ID NO: 325 is the determined cDNA sequence for clone 25311.
- SEQ ID NO: 326 is the determined cDNA sequence for clone 25310.
- SEQ ID NO: 327 is the determined cDNA sequence for clone 25302.
- SEQ ID NO: 328 is the determined cDNA sequence for clone 25315.
- 10 SEQ ID NO: 329 is the determined cDNA sequence for clone 25308.
- SEQ ID NO: 330 is the determined cDNA sequence for clone 25303.
- SEQ ID NO: 331-337 are the cDNA sequences of isoforms of the p53 tumor suppressor homologue, p63 (also referred to as L530S).
- SEQ ID NO: 338-344 are the amino acid sequences encoded by SEQ ID NO: 331-337,
- 15 respectively.
- SEQ ID NO: 345 is a second cDNA sequence for the antigen L763P.
- SEQ ID NO: 346 is the amino acid sequence encoded by the sequence of SEQ ID NO: 345.
- SEQ ID NO: 347 is a determined full-length cDNA sequence for L523S.
- SEQ ID NO: 348 is the predicted amino acid sequence encoded by SEQ ID NO: 347.
- 20 SEQ ID NO: 349 is the cDNA sequence encoding the N-terminal portion of L773P.
- SEQ ID NO: 350 is the amino acid sequence of the N-terminal portion of L773P.
- SEQ ID NO: 351 is polynucleotide sequence encoding the fusion of Ra12 and the N-terminal portion of L763P
- SEQ ID NO: 352 is the amino acid sequence of the fusion of Ra12 and the N-terminal
- 25 portion of L763P
- SEQ ID NO: 353 is polynucleotide sequence encoding the fusion of Ra12 and the C-terminal portion of L763P
- SEQ ID NO: 354 is the amino acid sequence of the fusion of Ra12 and the C-terminal portion of L763P

SEQ ID NO:355 is a primer.

SEQ ID NO:356 is a primer.

SEQ ID NO:357 is the protein sequence of expressed recombinant L762P.

SEQ ID NO:358 is the DNA sequence of expressed recombinant L762P.

5 SEQ ID NO:359 is a primer.

SEQ ID NO:360 is a primer.

SEQ ID NO:361 is the protein sequence of expressed recombinant L773P A.

SEQ ID NO:362 is the DNA sequence of expressed recombinant L773P A.

SEQ ID NO:363 is an epitope derived from clone L773P polypeptide.

10 SEQ ID NO:364 is a polynucleotide encoding the polypeptide of SEQ ID NO:363.

SEQ ID NO:365 is an epitope derived from clone L773P polypeptide.

SEQ ID NO:366 is a polynucleotide encoding the polypeptide of SEQ ID NO:365.

SEQ ID NO:367 is an epitope consisting of amino acids 571-590 of SEQ ID NO:161, clone L762.

15 SEQ ID NO:368 is the full-length DNA sequence for contig 13 (SEQ ID NO:125), also referred to as L761P.

SEQ ID NO:369 is the protein sequence encoded by the DNA sequence of SEQ ID NO:368.

SEQ ID NO:370 is an L762P DNA sequence from nucleotides 2071-2130.

20 SEQ ID NO:371 is an L762P DNA sequence from nucleotides 1441-1500.

SEQ ID NO:372 is an L762P DNA sequence from nucleotides 1936-1955.

SEQ ID NO:373 is an L762P DNA sequence from nucleotides 2620-2679.

SEQ ID NO:374 is an L762P DNA sequence from nucleotides 1801-1860.

SEQ ID NO:375 is an L762P DNA sequence from nucleotides 1531-1591.

25 SEQ ID NO:376 is the amino acid sequence of the L762P peptide encoded by SEQ ID NO: 373.

SEQ ID NO:377 is the amino acid sequence of the L762P peptide encoded by SEQ ID NO: 370.

SEQ ID NO:378 is the amino acid sequence of the L762P peptide encoded by SEQ ID NO: 372.

SEQ ID NO:379 is the amino acid sequence of the L762P peptide encoded by SEQ ID NO: 374.

5 SEQ ID NO:380 is the amino acid sequence of the L762P peptide encoded by SEQ ID NO: 371.

SEQ ID NO:381 is the amino acid sequence of the L762P peptide encoded by SEQ ID NO: 375.

Figure 1 shows the sequences of eleven L773P peptides.

10 Figure 2 shows that three CD4T cell lines (3C, 6G and 12B) recognized the appropriate L773P peptide as well as recombinant L773P and L773PA.

Figure 3 shows that individual CD4 T cell lines demonstrated cytokine release (IFN gamma) in response to the stimulating peptide but not the control peptide.

15 DETAILED DESCRIPTION OF THE INVENTION

As noted above, the present invention is generally directed to compositions and methods for the therapy and diagnosis of cancer, such as lung cancer. The compositions described herein may include lung tumor polypeptides, polynucleotides encoding such polypeptides, binding agents such as antibodies, antigen presenting cells (APCs) and/or immune system cells (*e.g.*, T cells). Polypeptides of the present invention generally comprise at least a portion (such as an immunogenic portion) of a lung tumor protein or a variant thereof. A "lung tumor protein" is a protein that is expressed in lung tumor cells at a level that is at least two fold, and preferably at least five fold, greater than the level of expression in a normal tissue, as determined using a representative assay provided herein. Certain lung tumor proteins are tumor proteins that react detectably (within an immunoassay, such as an ELISA or Western blot) with antisera of a patient afflicted with lung cancer. Polynucleotides of the subject invention generally comprise a DNA or RNA sequence that encodes all or a portion of such a polypeptide, or that is

20

25

complementary to such a sequence. Antibodies are generally immune system proteins, or antigen-binding fragments thereof, that are capable of binding to a polypeptide as described above. Antigen presenting cells include dendritic cells, macrophages, monocytes, fibroblasts and B-cells that express a polypeptide as described above. T cells that may be employed within such compositions are generally T cells that are specific for a polypeptide as described above.

The present invention is based on the discovery human lung tumor proteins. Sequences of polynucleotides encoding specific tumor proteins are provided in SEQ ID NO: 1-109, 111, 113, 115-151, 153, 154,157, 158, 160, 162-164, 167, 168, 171, 173, 175, 177-224, 255-337, 345, 347 and 349.

Therefore, in accordance with the above, and as described further below, the present invention provides illustrative polynucleotide compositions having sequences set forth in SEQ ID NO:1-109, 111, 113, 115-151, 153, 154,157, 158, 160, 162-164, 167, 168, 171, 173, 175, 177-224, 255-337, 345, 347 and 349, illustrative polypeptide compositions having amino acid sequences set forth in SEQ ID NO:110, 112, 114, 152, 155, 156, 159, 161, 165, 166, 169, 170, 172, 174, 176, 225, 252, 338-344, 346, 348, and 350, antibody compositions capable of binding such polypeptides, and numerous additional embodiments employing such compositions, for example in the detection, diagnosis and/or therapy of human lung cancer.

POLYNUCLEOTIDE COMPOSITIONS

As used herein, the terms "DNA segment" and "polynucleotide" refer to a DNA molecule that has been isolated free of total genomic DNA of a particular species. Therefore, a DNA segment encoding a polypeptide refers to a DNA segment that contains one or more coding sequences yet is substantially isolated away from, or purified free from, total genomic DNA of the species from which the DNA segment is obtained. Included within the terms "DNA segment" and "polynucleotide" are DNA segments and smaller fragments of such segments, and also recombinant vectors, including, for example, plasmids, cosmids, phagemids, phage, viruses, and the like.

As will be understood by those skilled in the art, the DNA segments of this invention can include genomic sequences, extra-genomic and plasmid-encoded sequences and smaller engineered gene segments that express, or may be adapted to express, proteins, polypeptides, peptides and the like. Such segments may be naturally isolated, or modified
 5 synthetically by the hand of man.

"Isolated," as used herein, means that a polynucleotide is substantially away from other coding sequences, and that the DNA segment does not contain large portions of unrelated coding DNA, such as large chromosomal fragments or other functional genes or polypeptide coding regions. Of course, this refers to the DNA segment as originally
 10 isolated, and does not exclude genes or coding regions later added to the segment by the hand of man.

As will be recognized by the skilled artisan, polynucleotides may be single-stranded (coding or antisense) or double-stranded, and may be DNA (genomic, cDNA or synthetic) or RNA molecules. RNA molecules include HnRNA molecules, which contain
 15 introns and correspond to a DNA molecule in a one-to-one manner, and mRNA molecules, which do not contain introns. Additional coding or non-coding sequences may, but need not, be present within a polynucleotide of the present invention, and a polynucleotide may, but need not, be linked to other molecules and/or support materials.

Polynucleotides may comprise a native sequence (*i.e.*, an endogenous
 20 sequence that encodes a lung tumor protein or a portion thereof) or may comprise a variant, or a biological or antigenic functional equivalent of such a sequence. Polynucleotide variants may contain one or more substitutions, additions, deletions and/or insertions, as further described below, preferably such that the immunogenicity of the encoded polypeptide is not diminished, relative to a native tumor protein. The effect on the
 25 immunogenicity of the encoded polypeptide may generally be assessed as described herein. The term "variants" also encompasses homologous genes of xenogenic origin.

When comparing polynucleotide or polypeptide sequences, two sequences are said to be "identical" if the sequence of nucleotides or amino acids in the two sequences is the same when aligned for maximum correspondence, as described below. Comparisons

between two sequences are typically performed by comparing the sequences over a comparison window to identify and compare local regions of sequence similarity. A “comparison window” as used herein, refers to a segment of at least about 20 contiguous positions, usually 30 to about 75, 40 to about 50, in which a sequence may be compared to a reference sequence of the same number of contiguous positions after the two sequences are optimally aligned.

Optimal alignment of sequences for comparison may be conducted using the Megalign program in the Lasergene suite of bioinformatics software (DNASTAR, Inc., Madison, WI), using default parameters. This program embodies several alignment schemes described in the following references: Dayhoff, M.O. (1978) A model of evolutionary change in proteins – Matrices for detecting distant relationships. In Dayhoff, M.O. (ed.) *Atlas of Protein Sequence and Structure*, National Biomedical Research Foundation, Washington DC Vol. 5, Suppl. 3, pp. 345-358; Hein J. (1990) Unified Approach to Alignment and Phylogenies pp. 626-645 *Methods in Enzymology* vol. 183, Academic Press, Inc., San Diego, CA; Higgins, D.G. and Sharp, P.M. (1989) *CABIOS* 5:151-153; Myers, E.W. and Muller W. (1988) *CABIOS* 4:11-17; Robinson, E.D. (1971) *Comb. Theor* 11:105; Santou, N. Nes, M. (1987) *Mol. Biol. Evol.* 4:406-425; Sneath, P.H.A. and Sokal, R.R. (1973) *Numerical Taxonomy – the Principles and Practice of Numerical Taxonomy*, Freeman Press, San Francisco, CA; Wilbur, W.J. and Lipman, D.J. (1983) *Proc. Natl. Acad., Sci. USA* 80:726-730.

Alternatively, optimal alignment of sequences for comparison may be conducted by the local identity algorithm of Smith and Waterman (1981) *Add. APL. Math* 2:482, by the identity alignment algorithm of Needleman and Wunsch (1970) *J. Mol. Biol.* 48:443, by the search for similarity methods of Pearson and Lipman (1988) *Proc. Natl. Acad. Sci. USA* 85: 2444, by computerized implementations of these algorithms (GAP, BESTFIT, BLAST, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group (GCG), 575 Science Dr., Madison, WI), or by inspection.

One preferred example of algorithms that are suitable for determining percent sequence identity and sequence similarity are the BLAST and BLAST 2.0

algorithms, which are described in Altschul *et al.* (1977) *Nucl. Acids Res.* 25:3389-3402 and Altschul *et al.* (1990) *J. Mol. Biol.* 215:403-410, respectively. BLAST and BLAST 2.0 can be used, for example with the parameters described herein, to determine percent sequence identity for the polynucleotides and polypeptides of the invention. Software for performing BLAST analyses is publicly available through the National Center for Biotechnology Information. In one illustrative example, cumulative scores can be calculated using, for nucleotide sequences, the parameters M (reward score for a pair of matching residues; always >0) and N (penalty score for mismatching residues; always <0). For amino acid sequences, a scoring matrix can be used to calculate the cumulative score.

Extension of the word hits in each direction are halted when: the cumulative alignment score falls off by the quantity X from its maximum achieved value; the cumulative score goes to zero or below, due to the accumulation of one or more negative-scoring residue alignments; or the end of either sequence is reached. The BLAST algorithm parameters W, T and X determine the sensitivity and speed of the alignment. The BLASTN program (for nucleotide sequences) uses as defaults a wordlength (W) of 11, and expectation (E) of 10, and the BLOSUM62 scoring matrix (see Henikoff and Henikoff (1989) *Proc. Natl. Acad. Sci. USA* 89:10915) alignments, (B) of 50, expectation (E) of 10, M=5, N=-4 and a comparison of both strands.

Preferably, the “percentage of sequence identity” is determined by comparing two optimally aligned sequences over a window of comparison of at least 20 positions, wherein the portion of the polynucleotide or polypeptide sequence in the comparison window may comprise additions or deletions (*i.e.*, gaps) of 20 percent or less, usually 5 to 15 percent, or 10 to 12 percent, as compared to the reference sequences (which does not comprise additions or deletions) for optimal alignment of the two sequences. The percentage is calculated by determining the number of positions at which the identical nucleic acid bases or amino acid residue occurs in both sequences to yield the number of matched positions, dividing the number of matched positions by the total number of positions in the reference sequence (*i.e.*, the window size) and multiplying the results by 100 to yield the percentage of sequence identity.

Therefore, the present invention encompasses polynucleotide and polypeptide sequences having substantial identity to the sequences disclosed herein, for example those comprising at least 50% sequence identity, preferably at least 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 96%, 97%, 98%, or 99% or higher, sequence
 5 identity compared to a polynucleotide or polypeptide sequence of this invention using the methods described herein, (*e.g.*, BLAST analysis using standard parameters, as described below). One skilled in this art will recognize that these values can be appropriately adjusted to determine corresponding identity of proteins encoded by two nucleotide
 10 positioning and the like.

In additional embodiments, the present invention provides isolated polynucleotides and polypeptides comprising various lengths of contiguous stretches of sequence identical to or complementary to one or more of the sequences disclosed herein. For example, polynucleotides are provided by this invention that comprise at least about
 15 15, 20, 30, 40, 50, 75, 100, 150, 200, 300, 400, 500 or 1000 or more contiguous nucleotides of one or more of the sequences disclosed herein as well as all intermediate lengths there between. It will be readily understood that "intermediate lengths", in this context, means any length between the quoted values, such as 16, 17, 18, 19, *etc.*; 21, 22, 23, *etc.*; 30, 31, 32, *etc.*; 50, 51, 52, 53, *etc.*; 100, 101, 102, 103, *etc.*; 150, 151, 152, 153, *etc.*; including all
 20 integers through 200-500; 500-1,000, and the like.

The polynucleotides of the present invention, or fragments thereof, regardless of the length of the coding sequence itself, may be combined with other DNA sequences, such as promoters, polyadenylation signals, additional restriction enzyme sites, multiple cloning sites, other coding segments, and the like, such that their overall length
 25 may vary considerably. It is therefore contemplated that a nucleic acid fragment of almost any length may be employed, with the total length preferably being limited by the ease of preparation and use in the intended recombinant DNA protocol. For example, illustrative DNA segments with total lengths of about 10,000, about 5000, about 3000, about 2,000, about 1,000, about 500, about 200, about 100, about 50 base pairs in length, and the like,

(including all intermediate lengths) are contemplated to be useful in many implementations of this invention.

In other embodiments, the present invention is directed to polynucleotides that are capable of hybridizing under moderately stringent conditions to a polynucleotide sequence provided herein, or a fragment thereof, or a complementary sequence thereof. Hybridization techniques are well known in the art of molecular biology. For purposes of illustration, suitable moderately stringent conditions for testing the hybridization of a polynucleotide of this invention with other polynucleotides include prewashing in a solution of 5 X SSC, 0.5% SDS, 1.0 mM EDTA (pH 8.0); hybridizing at 50°C-65°C, 5 X SSC, overnight; followed by washing twice at 65°C for 20 minutes with each of 2X, 0.5X and 0.2X SSC containing 0.1% SDS.

Moreover, it will be appreciated by those of ordinary skill in the art that, as a result of the degeneracy of the genetic code, there are many nucleotide sequences that encode a polypeptide as described herein. Some of these polynucleotides bear minimal homology to the nucleotide sequence of any native gene. Nonetheless, polynucleotides that vary due to differences in codon usage are specifically contemplated by the present invention. Further, alleles of the genes comprising the polynucleotide sequences provided herein are within the scope of the present invention. Alleles are endogenous genes that are altered as a result of one or more mutations, such as deletions, additions and/or substitutions of nucleotides. The resulting mRNA and protein may, but need not, have an altered structure or function. Alleles may be identified using standard techniques (such as hybridization, amplification and/or database sequence comparison).

Any polynucleotide that encodes a lung tumor protein or a portion or other variant thereof as described herein is encompassed by the present invention. Preferred polynucleotides comprise at least 15 consecutive nucleotides, preferably at least 30 consecutive nucleotides and more preferably at least 45 consecutive nucleotides, that encode a portion of a lung tumor protein. More preferably, a polynucleotide encodes an immunogenic portion of a lung tumor protein. Polynucleotides complementary to any such sequences are also encompassed by the present invention. Polynucleotides may be single-

stranded (coding or antisense) or double-stranded, and may be DNA (genomic, cDNA or synthetic) or RNA molecules. RNA molecules include HnRNA molecules, which contain introns and correspond to a DNA molecule in a one-to-one manner, and mRNA molecules, which do not contain introns. Additional coding or non-coding sequences may, but need not, be present within a polynucleotide of the present invention, and a polynucleotide may, but need not, be linked to other molecules and/or support materials.

PROBES AND PRIMERS

In other embodiments of the present invention, the polynucleotide sequences provided herein can be advantageously used as probes or primers for nucleic acid hybridization. As such, it is contemplated that nucleic acid segments that comprise a sequence region of at least about 15 nucleotide long contiguous sequence that has the same sequence as, or is complementary to, a 15 nucleotide long contiguous sequence disclosed herein will find particular utility. Longer contiguous identical or complementary sequences, *e.g.*, those of about 20, 30, 40, 50, 100, 200, 500, 1000 (including all intermediate lengths) and even up to full length sequences will also be of use in certain embodiments.

The ability of such nucleic acid probes to specifically hybridize to a sequence of interest will enable them to be of use in detecting the presence of complementary sequences in a given sample. However, other uses are also envisioned, such as the use of the sequence information for the preparation of mutant species primers, or primers for use in preparing other genetic constructions.

Polynucleotide molecules having sequence regions consisting of contiguous nucleotide stretches of 10-14, 15-20, 30, 50, or even of 100-200 nucleotides or so (including intermediate lengths as well), identical or complementary to a polynucleotide sequence disclosed herein, are particularly contemplated as hybridization probes for use in, *e.g.*, Southern and Northern blotting. This would allow a gene product, or fragment thereof, to be analyzed, both in diverse cell types and also in various bacterial cells. The total size of fragment, as well as the size of the complementary stretch(es), will ultimately

depend on the intended use or application of the particular nucleic acid segment. Smaller fragments will generally find use in hybridization embodiments, wherein the length of the contiguous complementary region may be varied, such as between about 15 and about 100 nucleotides, but larger contiguous complementarity stretches may be used, according to the length complementary sequences one wishes to detect.

The use of a hybridization probe of about 15-25 nucleotides in length allows the formation of a duplex molecule that is both stable and selective. Molecules having contiguous complementary sequences over stretches greater than 15 bases in length are generally preferred, though, in order to increase stability and selectivity of the hybrid, and thereby improve the quality and degree of specific hybrid molecules obtained. One will generally prefer to design nucleic acid molecules having gene-complementary stretches of 15 to 25 contiguous nucleotides, or even longer where desired.

Hybridization probes may be selected from any portion of any of the sequences disclosed herein. All that is required is to review the sequence set forth in SEQ ID NO:1-109, 111, 113, 115-151, 153, 154, 157, 158, 160, 162-164, 167, 168, 171, 173, 175, 177-224, 255-337, 345, 347 and 349, or to any continuous portion of the sequence, from about 15-25 nucleotides in length up to and including the full length sequence, that one wishes to utilize as a probe or primer. The choice of probe and primer sequences may be governed by various factors. For example, one may wish to employ primers from towards the termini of the total sequence.

Small polynucleotide segments or fragments may be readily prepared by, for example, directly synthesizing the fragment by chemical means, as is commonly practiced using an automated oligonucleotide synthesizer. Also, fragments may be obtained by application of nucleic acid reproduction technology, such as the PCR™ technology of U. S. Patent 4,683,202 (incorporated herein by reference), by introducing selected sequences into recombinant vectors for recombinant production, and by other recombinant DNA techniques generally known to those of skill in the art of molecular biology.

The nucleotide sequences of the invention may be used for their ability to selectively form duplex molecules with complementary stretches of the entire gene or gene

fragments of interest. Depending on the application envisioned, one will typically desire to employ varying conditions of hybridization to achieve varying degrees of selectivity of probe towards target sequence. For applications requiring high selectivity, one will typically desire to employ relatively stringent conditions to form the hybrids, *e.g.*, one will

5 select relatively low salt and/or high temperature conditions, such as provided by a salt concentration of from about 0.02 M to about 0.15 M salt at temperatures of from about 50°C to about 70°C. Such selective conditions tolerate little, if any, mismatch between the probe and the template or target strand, and would be particularly suitable for isolating related sequences.

10 Of course, for some applications, for example, where one desires to prepare mutants employing a mutant primer strand hybridized to an underlying template, less stringent (reduced stringency) hybridization conditions will typically be needed in order to allow formation of the heteroduplex. In these circumstances, one may desire to employ salt conditions such as those of from about 0.15 M to about 0.9 M salt, at temperatures ranging

15 from about 20°C to about 55°C. Cross-hybridizing species can thereby be readily identified as positively hybridizing signals with respect to control hybridizations. In any case, it is generally appreciated that conditions can be rendered more stringent by the addition of increasing amounts of formamide, which serves to destabilize the hybrid duplex in the same manner as increased temperature. Thus, hybridization conditions can be

20 readily manipulated, and thus will generally be a method of choice depending on the desired results.

POLYNUCLEOTIDE IDENTIFICATION AND CHARACTERIZATION

Polynucleotides may be identified, prepared and/or manipulated using any of a variety of well established techniques. For example, a polynucleotide may be

25 identified, as described in more detail below, by screening a microarray of cDNAs for tumor-associated expression (*i.e.*, expression that is at least two fold greater in a tumor than in normal tissue, as determined using a representative assay provided herein). Such screens may be performed, for example, using a Synteni microarray (Palo Alto, CA) according to

the manufacturer's instructions (and essentially as described by Schena *et al.*, *Proc. Natl. Acad. Sci. USA* 93:10614-10619, 1996 and Heller *et al.*, *Proc. Natl. Acad. Sci. USA* 94:2150-2155, 1997). Alternatively, polynucleotides may be amplified from cDNA prepared from cells expressing the proteins described herein, such as lung tumor cells.

- 5 Such polynucleotides may be amplified via polymerase chain reaction (PCR). For this approach, sequence-specific primers may be designed based on the sequences provided herein, and may be purchased or synthesized.

An amplified portion of a polynucleotide of the present invention may be used to isolate a full length gene from a suitable library (*e.g.*, a lung tumor cDNA library) using well known techniques. Within such techniques, a library (cDNA or genomic) is screened using one or more polynucleotide probes or primers suitable for amplification. Preferably, a library is size-selected to include larger molecules. Random primed libraries may also be preferred for identifying 5' and upstream regions of genes. Genomic libraries are preferred for obtaining introns and extending 5' sequences.

- 15 For hybridization techniques, a partial sequence may be labeled (*e.g.*, by nick-translation or end-labeling with ^{32}P) using well known techniques. A bacterial or bacteriophage library is then generally screened by hybridizing filters containing denatured bacterial colonies (or lawns containing phage plaques) with the labeled probe (*see* Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratories, Cold Spring Harbor, NY, 1989). Hybridizing colonies or plaques are selected and expanded, and the DNA is isolated for further analysis. cDNA clones may be analyzed to determine the amount of additional sequence by, for example, PCR using a primer from the partial sequence and a primer from the vector. Restriction maps and partial sequences may be generated to identify one or more overlapping clones. The complete sequence may then be determined using standard techniques, which may involve generating a series of deletion clones. The resulting overlapping sequences can then assembled into a single contiguous sequence. A full length cDNA molecule can be generated by ligating suitable fragments, using well known techniques.
- 25

Alternatively, there are numerous amplification techniques for obtaining a full length coding sequence from a partial cDNA sequence. Within such techniques, amplification is generally performed via PCR. Any of a variety of commercially available kits may be used to perform the amplification step. Primers may be designed using, for example, software well known in the art. Primers are preferably 22-30 nucleotides in length, have a GC content of at least 50% and anneal to the target sequence at temperatures of about 68°C to 72°C. The amplified region may be sequenced as described above, and overlapping sequences assembled into a contiguous sequence.

One such amplification technique is inverse PCR (*see* Triglia *et al.*, *Nucl. Acids Res.* 16:8186, 1988), which uses restriction enzymes to generate a fragment in the known region of the gene. The fragment is then circularized by intramolecular ligation and used as a template for PCR with divergent primers derived from the known region. Within an alternative approach, sequences adjacent to a partial sequence may be retrieved by amplification with a primer to a linker sequence and a primer specific to a known region. The amplified sequences are typically subjected to a second round of amplification with the same linker primer and a second primer specific to the known region. A variation on this procedure, which employs two primers that initiate extension in opposite directions from the known sequence, is described in WO 96/38591. Another such technique is known as "rapid amplification of cDNA ends" or RACE. This technique involves the use of an internal primer and an external primer, which hybridizes to a polyA region or vector sequence, to identify sequences that are 5' and 3' of a known sequence. Additional techniques include capture PCR (Lagerstrom *et al.*, *PCR Methods Applic.* 1:111-19, 1991) and walking PCR (Parker *et al.*, *Nucl. Acids. Res.* 19:3055-60, 1991). Other methods employing amplification may also be employed to obtain a full length cDNA sequence.

In certain instances, it is possible to obtain a full length cDNA sequence by analysis of sequences provided in an expressed sequence tag (EST) database, such as that available from GenBank. Searches for overlapping ESTs may generally be performed using well known programs (*e.g.*, NCBI BLAST searches), and such ESTs may be used to

generate a contiguous full length sequence. Full length DNA sequences may also be obtained by analysis of genomic fragments.

POLYNUCLEOTIDE EXPRESSION IN HOST CELLS

In other embodiments of the invention, polynucleotide sequences or
5 fragments thereof which encode polypeptides of the invention, or fusion proteins or functional equivalents thereof, may be used in recombinant DNA molecules to direct expression of a polypeptide in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences that encode substantially the same or a functionally equivalent amino acid sequence may be produced and these sequences may be used to
10 clone and express a given polypeptide.

As will be understood by those of skill in the art, it may be advantageous in some instances to produce polypeptide-encoding nucleotide sequences possessing non-naturally occurring codons. For example, codons preferred by a particular prokaryotic or eukaryotic host can be selected to increase the rate of protein expression or to produce a
15 recombinant RNA transcript having desirable properties, such as a half-life which is longer than that of a transcript generated from the naturally occurring sequence.

Moreover, the polynucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter polypeptide encoding sequences for a variety of reasons, including but not limited to, alterations which modify
20 the cloning, processing, and/or expression of the gene product. For example, DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. In addition, site-directed mutagenesis may be used to insert new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, or introduce mutations, and so
25 forth.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences may be ligated to a heterologous sequence to encode a fusion protein. For example, to screen peptide libraries for inhibitors of polypeptide activity, it

may be useful to encode a chimeric protein that can be recognized by a commercially available antibody. A fusion protein may also be engineered to contain a cleavage site located between the polypeptide-encoding sequence and the heterologous protein sequence, so that the polypeptide may be cleaved and purified away from the heterologous moiety.

5 Sequences encoding a desired polypeptide may be synthesized, in whole or in part, using chemical methods well known in the art (see Caruthers, M. H. *et al.* (1980) *Nucl. Acids Res. Symp. Ser.* 215-223, Horn, T. *et al.* (1980) *Nucl. Acids Res. Symp. Ser.* 225-232). Alternatively, the protein itself may be produced using chemical methods to synthesize the amino acid sequence of a polypeptide, or a portion thereof. For example,
10 peptide synthesis can be performed using various solid-phase techniques (Roberge, J. Y. *et al.* (1995) *Science* 269:202-204) and automated synthesis may be achieved, for example, using the ABI 431A Peptide Synthesizer (Perkin Elmer, Palo Alto, CA).

A newly synthesized peptide may be substantially purified by preparative high performance liquid chromatography (*e.g.*, Creighton, T. (1983) *Proteins, Structures and Molecular Principles*, WH Freeman and Co., New York, N.Y.) or other comparable
15 techniques available in the art. The composition of the synthetic peptides may be confirmed by amino acid analysis or sequencing (*e.g.*, the Edman degradation procedure). Additionally, the amino acid sequence of a polypeptide, or any part thereof, may be altered during direct synthesis and/or combined using chemical methods with sequences from
20 other proteins, or any part thereof, to produce a variant polypeptide.

In order to express a desired polypeptide, the nucleotide sequences encoding the polypeptide, or functional equivalents, may be inserted into appropriate expression vector, *i.e.*, a vector which contains the necessary elements for the transcription and translation of the inserted coding sequence. Methods which are well known to those
25 skilled in the art may be used to construct expression vectors containing sequences encoding a polypeptide of interest and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. Such techniques are described in Sambrook, J. *et al.* (1989) *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Press,

Plainview, N.Y., and Ausubel, F. M. *et al.* (1989) Current Protocols in Molecular Biology, John Wiley & Sons, New York. N.Y.

A variety of expression vector/host systems may be utilized to contain and express polynucleotide sequences. These include, but are not limited to, microorganisms
 5 such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with virus expression vectors (*e.g.*, baculovirus); plant cell systems transformed with virus expression vectors (*e.g.*, cauliflower mosaic virus, CaMV; tobacco mosaic virus, TMV) or with bacterial expression vectors (*e.g.*, Ti or pBR322 plasmids); or animal cell
 10 systems.

The "control elements" or "regulatory sequences" present in an expression vector are those non-translated regions of the vector--enhancers, promoters, 5' and 3' untranslated regions--which interact with host cellular proteins to carry out transcription and translation. Such elements may vary in their strength and specificity. Depending on the
 15 vector system and host utilized, any number of suitable transcription and translation elements, including constitutive and inducible promoters, may be used. For example, when cloning in bacterial systems, inducible promoters such as the hybrid lacZ promoter of the PBLUESCRIPT phagemid (Stratagene, La Jolla, Calif.) or PSPORT1 plasmid (Gibco BRL, Gaithersburg, MD) and the like may be used. In mammalian cell systems, promoters from
 20 mammalian genes or from mammalian viruses are generally preferred. If it is necessary to generate a cell line that contains multiple copies of the sequence encoding a polypeptide, vectors based on SV40 or EBV may be advantageously used with an appropriate selectable marker.

In bacterial systems, a number of expression vectors may be selected
 25 depending upon the use intended for the expressed polypeptide. For example, when large quantities are needed, for example for the induction of antibodies, vectors which direct high level expression of fusion proteins that are readily purified may be used. Such vectors include, but are not limited to, the multifunctional *E. coli* cloning and expression vectors such as BLUESCRIPT (Stratagene), in which the sequence encoding the polypeptide of

interest may be ligated into the vector in frame with sequences for the amino-terminal Met and the subsequent 7 residues of β -galactosidase so that a hybrid protein is produced; pIN vectors (Van Heeke, G. and S. M. Schuster (1989) *J. Biol. Chem.* 264:5503-5509); and the like. pGEX Vectors (Promega, Madison, Wis.) may also be used to express foreign polypeptides as fusion proteins with glutathione S-transferase (GST). In general, such fusion proteins are soluble and can easily be purified from lysed cells by adsorption to glutathione-agarose beads followed by elution in the presence of free glutathione. Proteins made in such systems may be designed to include heparin, thrombin, or factor XA protease cleavage sites so that the cloned polypeptide of interest can be released from the GST moiety at will.

In the yeast, *Saccharomyces cerevisiae*, a number of vectors containing constitutive or inducible promoters such as alpha factor, alcohol oxidase, and PGH may be used. For reviews, see Ausubel *et al.* (supra) and Grant *et al.* (1987) *Methods Enzymol.* 153:516-544.

In cases where plant expression vectors are used, the expression of sequences encoding polypeptides may be driven by any of a number of promoters. For example, viral promoters such as the 35S and 19S promoters of CaMV may be used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) *EMBO J.* 6:307-311. Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used (Coruzzi, G. *et al.* (1984) *EMBO J.* 3:1671-1680; Broglie, R. *et al.* (1984) *Science* 224:838-843; and Winter, J. *et al.* (1991) *Results Probl. Cell Differ.* 17:85-105). These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. Such techniques are described in a number of generally available reviews (see, for example, Hobbs, S. or Murry, L. E. in McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York, N.Y.; pp. 191-196).

An insect system may also be used to express a polypeptide of interest. For example, in one such system, *Autographa californica* nuclear polyhedrosis virus (AcNPV) is used as a vector to express foreign genes in *Spodoptera frugiperda* cells or in

Trichoplusia larvae. The sequences encoding the polypeptide may be cloned into a non-essential region of the virus, such as the polyhedrin gene, and placed under control of the polyhedrin promoter. Successful insertion of the polypeptide-encoding sequence will render the polyhedrin gene inactive and produce recombinant virus lacking coat protein.

- 5 The recombinant viruses may then be used to infect, for example, *S. frugiperda* cells or *Trichoplusia* larvae in which the polypeptide of interest may be expressed (Engelhard, E. K. *et al.* (1994) *Proc. Natl. Acad. Sci.* 91 :3224-3227).

In mammalian host cells, a number of viral-based expression systems are generally available. For example, in cases where an adenovirus is used as an expression
 10 vector, sequences encoding a polypeptide of interest may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain a viable virus which is capable of expressing the polypeptide in infected host cells (Logan, J. and Shenk, T. (1984) *Proc. Natl. Acad. Sci.* 81:3655-3659). In addition,
 15 transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells.

Specific initiation signals may also be used to achieve more efficient translation of sequences encoding a polypeptide of interest. Such signals include the ATG initiation codon and adjacent sequences. In cases where sequences encoding the
 20 polypeptide, its initiation codon, and upstream sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a portion thereof, is inserted, exogenous translational control signals including the ATG initiation codon should be provided. Furthermore, the initiation codon should be in the correct reading frame to ensure
 25 translation of the entire insert. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers which are appropriate for the particular cell system which is used, such as those described in the literature (Scharf, D. *et al.* (1994) *Results Probl. Cell Differ.* 20:125-162).

In addition, a host cell strain may be chosen for its ability to modulate the expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" form of the protein may also be used to facilitate correct insertion, folding and/or function. Different host cells such as CHO, HeLa, MDCK, HEK293, and WI38, which have specific cellular machinery and characteristic mechanisms for such post-translational activities, may be chosen to ensure the correct modification and processing of the foreign protein.

For long-term, high-yield production of recombinant proteins, stable expression is generally preferred. For example, cell lines which stably express a polynucleotide of interest may be transformed using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for 1-2 days in an enriched media before they are switched to selective media. The purpose of the selectable marker is to confer resistance to selection, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be proliferated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase (Wigler, M. *et al.* (1977) *Cell* 11:223-32) and adenine phosphoribosyltransferase (Lowy, I. *et al.* (1990) *Cell* 22:817-23) genes which can be employed in tk.sup.- or aprt.sup.- cells, respectively. Also, antimetabolite, antibiotic or herbicide resistance can be used as the basis for selection; for example, dhfr which confers resistance to methotrexate (Wigler, M. *et al.* (1980) *Proc. Natl. Acad. Sci.* 77:3567-70); npt, which confers resistance to the aminoglycosides, neomycin and G-418 (Colbere-Garapin, F. *et al.* (1981) *J. Mol. Biol.* 150:1-14); and als or pat, which confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively (Murry, *supra*). Additional selectable genes have been

described, for example, *trpB*, which allows cells to utilize indole in place of tryptophan, or *hisD*, which allows cells to utilize histinol in place of histidine (Hartman, S. C. and R. C. Mulligan (1988) *Proc. Natl. Acad. Sci.* 85:8047-51). Recently, the use of visible markers has gained popularity with such markers as anthocyanins, beta-glucuronidase and its substrate GUS, and luciferase and its substrate luciferin, being widely used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system (Rhodes, C. A. *et al.* (1995) *Methods Mol. Biol.* 55:121-131).

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, its presence and expression may need to be confirmed. For example, if the sequence encoding a polypeptide is inserted within a marker gene sequence, recombinant cells containing sequences can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a polypeptide-encoding sequence under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

Alternatively, host cells which contain and express a desired polynucleotide sequence may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations and protein bioassay or immunoassay techniques which include membrane, solution, or chip based technologies for the detection and/or quantification of nucleic acid or protein.

A variety of protocols for detecting and measuring the expression of polynucleotide-encoded products, using either polyclonal or monoclonal antibodies specific for the product are known in the art. Examples include enzyme-linked immunosorbent assay (ELISA), radioimmunoassay (RIA), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on a given polypeptide may be preferred for some applications, but a competitive binding assay may also be employed. These and other assays are described, among other places, in Hampton, R. *et al.* (1990; Serological

Methods, a Laboratory Manual, APS Press, St Paul. Minn.) and Maddox, D. E. *et al.* (1983; *J. Exp. Med.* 158:1211-1216).

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides include oligolabeling, nick translation, end-labeling or PCR amplification using a labeled nucleotide. Alternatively, the sequences, or any portions thereof may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits. Suitable reporter molecules or labels, which may be used include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with a polynucleotide sequence of interest may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a recombinant cell may be secreted or contained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides of the invention may be designed to contain signal sequences which direct secretion of the encoded polypeptide through a prokaryotic or eukaryotic cell membrane. Other recombinant constructions may be used to join sequences encoding a polypeptide of interest to nucleotide sequence encoding a polypeptide domain which will facilitate purification of soluble proteins. Such purification facilitating domains include, but are not limited to, metal chelating peptides such as histidine-tryptophan modules that allow purification on immobilized metals, protein A domains that allow purification on immobilized immunoglobulin, and the domain utilized in the FLAGS extension/affinity purification system (Immunex Corp., Seattle, Wash.). The inclusion of cleavable linker sequences such as those specific for Factor XA or enterokinase (Invitrogen. San Diego, Calif.) between the purification domain and the

encoded polypeptide may be used to facilitate purification. One such expression vector provides for expression of a fusion protein containing a polypeptide of interest and a nucleic acid encoding 6 histidine residues preceding a thioredoxin or an enterokinase cleavage site. The histidine residues facilitate purification on IMIAC (immobilized metal ion affinity chromatography) as described in Porath, J. *et al.* (1992, *Prot. Exp. Purif.* 3:263-281) while the enterokinase cleavage site provides a means for purifying the desired polypeptide from the fusion protein. A discussion of vectors which contain fusion proteins is provided in Kroll, D. J. *et al.* (1993; *DNA Cell Biol.* 12:441-453).

In addition to recombinant production methods, polypeptides of the invention, and fragments thereof, may be produced by direct peptide synthesis using solid-phase techniques (Merrifield J. (1963) *J. Am. Chem. Soc.* 85:2149-2154). Protein synthesis may be performed using manual techniques or by automation. Automated synthesis may be achieved, for example, using Applied Biosystems 431A Peptide Synthesizer (Perkin Elmer). Alternatively, various fragments may be chemically synthesized separately and combined using chemical methods to produce the full length molecule.

SITE-SPECIFIC MUTAGENESIS

Site-specific mutagenesis is a technique useful in the preparation of individual peptides, or biologically functional equivalent polypeptides, through specific mutagenesis of the underlying polynucleotides that encode them. The technique, well-known to those of skill in the art, further provides a ready ability to prepare and test sequence variants, for example, incorporating one or more of the foregoing considerations, by introducing one or more nucleotide sequence changes into the DNA. Site-specific mutagenesis allows the production of mutants through the use of specific oligonucleotide sequences which encode the DNA sequence of the desired mutation, as well as a sufficient number of adjacent nucleotides, to provide a primer sequence of sufficient size and sequence complexity to form a stable duplex on both sides of the deletion junction being traversed. Mutations may be employed in a selected polynucleotide sequence to improve, alter, decrease, modify, or otherwise change the properties of the polynucleotide itself,

and/or alter the properties, activity, composition, stability, or primary sequence of the encoded polypeptide.

In certain embodiments of the present invention, the inventors contemplate the mutagenesis of the disclosed polynucleotide sequences to alter one or more properties of the encoded polypeptide, such as the antigenicity of a polypeptide vaccine. The techniques of site-specific mutagenesis are well-known in the art, and are widely used to create variants of both polypeptides and polynucleotides. For example, site-specific mutagenesis is often used to alter a specific portion of a DNA molecule. In such embodiments, a primer comprising typically about 14 to about 25 nucleotides or so in length is employed, with about 5 to about 10 residues on both sides of the junction of the sequence being altered.

As will be appreciated by those of skill in the art, site-specific mutagenesis techniques have often employed a phage vector that exists in both a single stranded and double stranded form. Typical vectors useful in site-directed mutagenesis include vectors such as the M13 phage. These phage are readily commercially-available and their use is generally well-known to those skilled in the art. Double-stranded plasmids are also routinely employed in site directed mutagenesis that eliminates the step of transferring the gene of interest from a plasmid to a phage.

In general, site-directed mutagenesis in accordance herewith is performed by first obtaining a single-stranded vector or melting apart of two strands of a double-stranded vector that includes within its sequence a DNA sequence that encodes the desired peptide. An oligonucleotide primer bearing the desired mutated sequence is prepared, generally synthetically. This primer is then annealed with the single-stranded vector, and subjected to DNA polymerizing enzymes such as *E. coli* polymerase I Klenow fragment, in order to complete the synthesis of the mutation-bearing strand. Thus, a heteroduplex is formed wherein one strand encodes the original non-mutated sequence and the second strand bears the desired mutation. This heteroduplex vector is then used to transform appropriate cells, such as *E. coli* cells, and clones are selected which include recombinant vectors bearing the mutated sequence arrangement.

The preparation of sequence variants of the selected peptide-encoding DNA segments using site-directed mutagenesis provides a means of producing potentially useful species and is not meant to be limiting as there are other ways in which sequence variants of peptides and the DNA sequences encoding them may be obtained. For example, recombinant vectors encoding the desired peptide sequence may be treated with mutagenic agents, such as hydroxylamine, to obtain sequence variants. Specific details regarding these methods and protocols are found in the teachings of Maloy *et al.*, 1994; Segal, 1976; Prokop and Bajpai, 1991; Kuby, 1994; and Maniatis *et al.*, 1982, each incorporated herein by reference, for that purpose.

As used herein, the term “oligonucleotide directed mutagenesis procedure” refers to template-dependent processes and vector-mediated propagation which result in an increase in the concentration of a specific nucleic acid molecule relative to its initial concentration, or in an increase in the concentration of a detectable signal, such as amplification. As used herein, the term “oligonucleotide directed mutagenesis procedure” is intended to refer to a process that involves the template-dependent extension of a primer molecule. The term template dependent process refers to nucleic acid synthesis of an RNA or a DNA molecule wherein the sequence of the newly synthesized strand of nucleic acid is dictated by the well-known rules of complementary base pairing (see, for example, Watson, 1987). Typically, vector mediated methodologies involve the introduction of the nucleic acid fragment into a DNA or RNA vector, the clonal amplification of the vector, and the recovery of the amplified nucleic acid fragment. Examples of such methodologies are provided by U. S. Patent No. 4,237,224, specifically incorporated herein by reference in its entirety.

POLYNUCLEOTIDE AMPLIFICATION TECHNIQUES

A number of template dependent processes are available to amplify the target sequences of interest present in a sample. One of the best known amplification methods is the polymerase chain reaction (PCRTM) which is described in detail in U.S. Patent Nos. 4,683,195, 4,683,202 and 4,800,159, each of which is incorporated herein by

reference in its entirety. Briefly, in PCRTM, two primer sequences are prepared which are complementary to regions on opposite complementary strands of the target sequence. An excess of deoxynucleoside triphosphates is added to a reaction mixture along with a DNA polymerase (*e.g.*, *Taq* polymerase). If the target sequence is present in a sample, the

5 primers will bind to the target and the polymerase will cause the primers to be extended along the target sequence by adding on nucleotides. By raising and lowering the temperature of the reaction mixture, the extended primers will dissociate from the target to form reaction products, excess primers will bind to the target and to the reaction product and the process is repeated. Preferably reverse transcription and PCRTM amplification

10 procedure may be performed in order to quantify the amount of mRNA amplified. Polymerase chain reaction methodologies are well known in the art.

Another method for amplification is the ligase chain reaction (referred to as LCR), disclosed in Eur. Pat. Appl. Publ. No. 320,308 (specifically incorporated herein by reference in its entirety). In LCR, two complementary probe pairs are prepared, and in the

15 presence of the target sequence, each pair will bind to opposite complementary strands of the target such that they abut. In the presence of a ligase, the two probe pairs will link to form a single unit. By temperature cycling, as in PCRTM, bound ligated units dissociate from the target and then serve as "target sequences" for ligation of excess probe pairs. U.S. Patent No. 4,883,750, incorporated herein by reference in its entirety, describes an

20 alternative method of amplification similar to LCR for binding probe pairs to a target sequence.

Qbeta Replicase, described in PCT Intl. Pat. Appl. Publ. No. PCT/US87/00880, incorporated herein by reference in its entirety, may also be used as still another amplification method in the present invention. In this method, a replicative

25 sequence of RNA that has a region complementary to that of a target is added to a sample in the presence of an RNA polymerase. The polymerase will copy the replicative sequence that can then be detected.

An isothermal amplification method, in which restriction endonucleases and ligases are used to achieve the amplification of target molecules that contain nucleotide

5'-[α -thio]triphosphates in one strand of a restriction site (Walker *et al.*, 1992, incorporated herein by reference in its entirety), may also be useful in the amplification of nucleic acids in the present invention.

Strand Displacement Amplification (SDA) is another method of carrying out isothermal amplification of nucleic acids which involves multiple rounds of strand displacement and synthesis, *i.e.* nick translation. A similar method, called Repair Chain Reaction (RCR) is another method of amplification which may be useful in the present invention and is involves annealing several probes throughout a region targeted for amplification, followed by a repair reaction in which only two of the four bases are present. The other two bases can be added as biotinylated derivatives for easy detection. A similar approach is used in SDA.

Sequences can also be detected using a cyclic probe reaction (CPR). In CPR, a probe having a 3' and 5' sequences of non-target DNA and an internal or "middle" sequence of the target protein specific RNA is hybridized to DNA which is present in a sample. Upon hybridization, the reaction is treated with RNaseH, and the products of the probe are identified as distinctive products by generating a signal that is released after digestion. The original template is annealed to another cycling probe and the reaction is repeated. Thus, CPR involves amplifying a signal generated by hybridization of a probe to a target gene specific expressed nucleic acid.

Still other amplification methods described in Great Britain Pat. Appl. No. 2 202 328, and in PCT Intl. Pat. Appl. Publ. No. PCT/US89/01025, each of which is incorporated herein by reference in its entirety, may be used in accordance with the present invention. In the former application, "modified" primers are used in a PCR-like, template and enzyme dependent synthesis. The primers may be modified by labeling with a capture moiety (*e.g.*, biotin) and/or a detector moiety (*e.g.*, enzyme). In the latter application, an excess of labeled probes is added to a sample. In the presence of the target sequence, the probe binds and is cleaved catalytically. After cleavage, the target sequence is released intact to be bound by excess probe. Cleavage of the labeled probe signals the presence of the target sequence.

Other nucleic acid amplification procedures include transcription-based amplification systems (TAS) (Kwoh *et al.*, 1989; PCT Intl. Pat. Appl. Publ. No. WO 88/10315, incorporated herein by reference in its entirety), including nucleic acid sequence based amplification (NASBA) and 3SR. In NASBA, the nucleic acids can be prepared for amplification by standard phenol/chloroform extraction, heat denaturation of a sample, treatment with lysis buffer and minispin columns for isolation of DNA and RNA or guanidinium chloride extraction of RNA. These amplification techniques involve annealing a primer that has sequences specific to the target sequence. Following polymerization, DNA/RNA hybrids are digested with RNase H while double stranded DNA molecules are heat-denatured again. In either case the single stranded DNA is made fully double stranded by addition of second target-specific primer, followed by polymerization. The double stranded DNA molecules are then multiply transcribed by a polymerase such as T7 or SP6. In an isothermal cyclic reaction, the RNAs are reverse transcribed into DNA, and transcribed once again with a polymerase such as T7 or SP6.

15 The resulting products, whether truncated or complete, indicate target-specific sequences.

Eur. Pat. Appl. Publ. No. 329,822, incorporated herein by reference in its entirety, disclose a nucleic acid amplification process involving cyclically synthesizing single-stranded RNA ("ssRNA"), ssDNA, and double-stranded DNA (dsDNA), which may be used in accordance with the present invention. The ssRNA is a first template for a first primer oligonucleotide, which is elongated by reverse transcriptase (RNA-dependent DNA polymerase). The RNA is then removed from resulting DNA:RNA duplex by the action of ribonuclease H (RNase H, an RNase specific for RNA in a duplex with either DNA or RNA). The resultant ssDNA is a second template for a second primer, which also includes the sequences of an RNA polymerase promoter (exemplified by T7 RNA polymerase) 5' to its homology to its template. This primer is then extended by DNA polymerase (exemplified by the large "Klenow" fragment of *E. coli* DNA polymerase I), resulting as a double-stranded DNA ("dsDNA") molecule, having a sequence identical to that of the original RNA between the primers and having additionally, at one end, a promoter sequence. This promoter sequence can be used by the appropriate RNA polymerase to

make many RNA copies of the DNA. These copies can then re-enter the cycle leading to very swift amplification. With proper choice of enzymes, this amplification can be done isothermally without addition of enzymes at each cycle. Because of the cyclical nature of this process, the starting sequence can be chosen to be in the form of either DNA or RNA.

5 PCT Intl. Pat. Appl. Publ. No. WO 89/06700, incorporated herein by reference in its entirety, disclose a nucleic acid sequence amplification scheme based on the hybridization of a promoter/primer sequence to a target single-stranded DNA ("ssDNA") followed by transcription of many RNA copies of the sequence. This scheme is not cyclic; *i.e.* new templates are not produced from the resultant RNA transcripts. Other
10 amplification methods include "RACE" (Frohman, 1990), and "one-sided PCR" (Ohara, 1989) which are well-known to those of skill in the art.

Methods based on ligation of two (or more) oligonucleotides in the presence of nucleic acid having the sequence of the resulting "di-oligonucleotide", thereby amplifying the di-oligonucleotide (Wu and Dean, 1996, incorporated herein by reference in
15 its entirety), may also be used in the amplification of DNA sequences of the present invention.

BIOLOGICAL FUNCTIONAL EQUIVALENTS

Modification and changes may be made in the structure of the polynucleotides and polypeptides of the present invention and still obtain a functional
20 molecule that encodes a polypeptide with desirable characteristics. As mentioned above, it is often desirable to introduce one or more mutations into a specific polynucleotide sequence. In certain circumstances, the resulting encoded polypeptide sequence is altered by this mutation, or in other cases, the sequence of the polypeptide is unchanged by one or more mutations in the encoding polynucleotide.

25 When it is desirable to alter the amino acid sequence of a polypeptide to create an equivalent, or even an improved, second-generation molecule, the amino acid changes may be achieved by changing one or more of the codons of the encoding DNA sequence, according to Table 1.

For example, certain amino acids may be substituted for other amino acids in a protein structure without appreciable loss of interactive binding capacity with structures such as, for example, antigen-binding regions of antibodies or binding sites on substrate molecules. Since it is the interactive capacity and nature of a protein that defines that protein's biological functional activity, certain amino acid sequence substitutions can be made in a protein sequence, and, of course, its underlying DNA coding sequence, and nevertheless obtain a protein with like properties. It is thus contemplated by the inventors that various changes may be made in the peptide sequences of the disclosed compositions, or corresponding DNA sequences which encode said peptides without appreciable loss of their biological utility or activity.

TABLE 1

Amino Acids			Codons					
Alanine	Ala	A	GCA	GCC	GCG	GCU		
Cysteine	Cys	C	UGC	UGU				
Aspartic acid	Asp	D	GAC	GAU				
Glutamic acid	Glu	E	GAA	GAG				
Phenylalanine	Phe	F	UUC	UUU				
Glycine	Gly	G	GGA	GGC	GGG	GGU		
Histidine	His	H	CAC	CAU				
Isoleucine	Ile	I	AUA	AUC	AUU			
Lysine	Lys	K	AAA	AAG				
Leucine	Leu	L	UUA	UUG	CUA	CUC	CUG	CUU
Methionine	Met	M	AUG					
Asparagine	Asn	N	AAC	AAU				
Proline	Pro	P	CCA	CCC	CCG	CCU		
Glutamine	Gln	Q	CAA	CAG				
Arginine	Arg	R	AGA	AGG	CGA	CGC	CGG	CGU
Serine	Ser	S	AGC	AGU	UCA	UCC	UCG	UCU
Threonine	Thr	T	ACA	ACC	ACG	ACU		
Valine	Val	V	GUA	GUC	GUG	GUU		
Tryptophan	Trp	W	UGG					
Tyrosine	Tyr	Y	UAC	UAU				

In making such changes, the hydropathic index of amino acids may be considered. The importance of the hydropathic amino acid index in conferring interactive
5 biologic function on a protein is generally understood in the art (Kyte and Doolittle, 1982, incorporated herein by reference). It is accepted that the relative hydropathic character of the amino acid contributes to the secondary structure of the resultant protein, which in turn defines the interaction of the protein with other molecules, for example, enzymes, substrates, receptors, DNA, antibodies, antigens, and the like. Each amino acid has been

assigned a hydropathic index on the basis of its hydrophobicity and charge characteristics (Kyte and Doolittle, 1982). These values are: isoleucine (+4.5); valine (+4.2); leucine (+3.8); phenylalanine (+2.8); cysteine/cystine (+2.5); methionine (+1.9); alanine (+1.8); glycine (−0.4); threonine (−0.7); serine (−0.8); tryptophan (−0.9); tyrosine (−1.3); proline (−1.6); histidine (−3.2); glutamate (−3.5); glutamine (−3.5); aspartate (−3.5); asparagine (−3.5); lysine (−3.9); and arginine (−4.5).

It is known in the art that certain amino acids may be substituted by other amino acids having a similar hydropathic index or score and still result in a protein with similar biological activity, *i.e.* still obtain a biological functionally equivalent protein. In making such changes, the substitution of amino acids whose hydropathic indices are within ± 2 is preferred, those within ± 1 are particularly preferred, and those within ± 0.5 are even more particularly preferred. It is also understood in the art that the substitution of like amino acids can be made effectively on the basis of hydrophilicity. U. S. Patent 4,554,101 (specifically incorporated herein by reference in its entirety), states that the greatest local average hydrophilicity of a protein, as governed by the hydrophilicity of its adjacent amino acids, correlates with a biological property of the protein.

As detailed in U. S. Patent 4,554,101, the following hydrophilicity values have been assigned to amino acid residues: arginine (+3.0); lysine (+3.0); aspartate (+3.0 \pm 1); glutamate (+3.0 \pm 1); serine (+0.3); asparagine (+0.2); glutamine (+0.2); glycine (0); threonine (−0.4); proline (−0.5 \pm 1); alanine (−0.5); histidine (−0.5); cysteine (−1.0); methionine (−1.3); valine (−1.5); leucine (−1.8); isoleucine (−1.8); tyrosine (−2.3); phenylalanine (−2.5); tryptophan (−3.4). It is understood that an amino acid can be substituted for another having a similar hydrophilicity value and still obtain a biologically equivalent, and in particular, an immunologically equivalent protein. In such changes, the substitution of amino acids whose hydrophilicity values are within ± 2 is preferred, those within ± 1 are particularly preferred, and those within ± 0.5 are even more particularly preferred.

As outlined above, amino acid substitutions are generally therefore based on the relative similarity of the amino acid side-chain substituents, for example, their

hydrophobicity, hydrophilicity, charge, size, and the like. Exemplary substitutions that take various of the foregoing characteristics into consideration are well known to those of skill in the art and include: arginine and lysine; glutamate and aspartate; serine and threonine; glutamine and asparagine; and valine, leucine and isoleucine.

5 In addition, any polynucleotide may be further modified to increase stability *in vivo*. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends; the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages in the backbone; and/or the inclusion of nontraditional bases such as inosine, queosine and wybutosine, as well as acetyl- methyl-, thio- and other
10 modified forms of adenine, cytidine, guanine, thymine and uridine.

IN VIVO POLYNUCLEOTIDE DELIVERY TECHNIQUES

In additional embodiments, genetic constructs comprising one or more of the polynucleotides of the invention are introduced into cells *in vivo*. This may be achieved using any of a variety of well known approaches, several of which are outlined below for
15 the purpose of illustration.

1. ADENOVIRUS

One of the preferred methods for *in vivo* delivery of one or more nucleic acid sequences involves the use of an adenovirus expression vector. "Adenovirus expression vector" is meant to include those constructs containing adenovirus sequences
20 sufficient to (a) support packaging of the construct and (b) to express a polynucleotide that has been cloned therein in a sense or antisense orientation. Of course, in the context of an antisense construct, expression does not require that the gene product be synthesized.

The expression vector comprises a genetically engineered form of an adenovirus. Knowledge of the genetic organization of adenovirus, a 36 kb, linear, double-
25 stranded DNA virus, allows substitution of large pieces of adenoviral DNA with foreign sequences up to 7 kb (Grunhaus and Horwitz, 1992). In contrast to retrovirus, the adenoviral infection of host cells does not result in chromosomal integration because

adenoviral DNA can replicate in an episomal manner without potential genotoxicity. Also, adenoviruses are structurally stable, and no genome rearrangement has been detected after extensive amplification. Adenovirus can infect virtually all epithelial cells regardless of their cell cycle stage. So far, adenoviral infection appears to be linked only to mild disease
 5 such as acute respiratory disease in humans.

Adenovirus is particularly suitable for use as a gene transfer vector because of its mid-sized genome, ease of manipulation, high titer, wide target-cell range and high infectivity. Both ends of the viral genome contain 100-200 base pair inverted repeats (ITRs), which are *cis* elements necessary for viral DNA replication and packaging. The
 10 early (E) and late (L) regions of the genome contain different transcription units that are divided by the onset of viral DNA replication. The E1 region (E1A and E1B) encodes proteins responsible for the regulation of transcription of the viral genome and a few cellular genes. The expression of the E2 region (E2A and E2B) results in the synthesis of the proteins for viral DNA replication. These proteins are involved in DNA replication,
 15 late gene expression and host cell shut-off (Renan, 1990). The products of the late genes, including the majority of the viral capsid proteins, are expressed only after significant processing of a single primary transcript issued by the major late promoter (MLP). The MLP, (located at 16.8 m.u.) is particularly efficient during the late phase of infection, and all the mRNA's issued from this promoter possess a 5'-tripartite leader (TPL) sequence
 20 which makes them preferred mRNA's for translation.

In a current system, recombinant adenovirus is generated from homologous recombination between shuttle vector and provirus vector. Due to the possible recombination between two proviral vectors, wild-type adenovirus may be generated from this process. Therefore, it is critical to isolate a single clone of virus from an individual
 25 plaque and examine its genomic structure.

Generation and propagation of the current adenovirus vectors, which are replication deficient, depend on a unique helper cell line, designated 293, which was transformed from human embryonic kidney cells by Ad5 DNA fragments and constitutively expresses E1 proteins (Graham *et al.*, 1977). Since the E3 region is

dispensable from the adenovirus genome (Jones and Shenk, 1978), the current adenovirus vectors, with the help of 293 cells, carry foreign DNA in either the E1, the D3 or both regions (Graham and Prevec, 1991). In nature, adenovirus can package approximately 105% of the wild-type genome (Ghosh-Choudhury *et al.*, 1987), providing capacity for about 2 extra kB of DNA. Combined with the approximately 5.5 kB of DNA that is replaceable in the E1 and E3 regions, the maximum capacity of the current adenovirus vector is under 7.5 kB, or about 15% of the total length of the vector. More than 80% of the adenovirus viral genome remains in the vector backbone and is the source of vector-borne cytotoxicity. Also, the replication deficiency of the E1-deleted virus is incomplete.

For example, leakage of viral gene expression has been observed with the currently available vectors at high multiplicities of infection (MOI) (Mulligan, 1993).

Helper cell lines may be derived from human cells such as human embryonic kidney cells, muscle cells, hematopoietic cells or other human embryonic mesenchymal or epithelial cells. Alternatively, the helper cells may be derived from the cells of other mammalian species that are permissive for human adenovirus. Such cells include, *e.g.*, Vero cells or other monkey embryonic mesenchymal or epithelial cells. As stated above, the currently preferred helper cell line is 293.

Recently, Racher *et al.* (1995) disclosed improved methods for culturing 293 cells and propagating adenovirus. In one format, natural cell aggregates are grown by inoculating individual cells into 1 liter siliconized spinner flasks (Techne, Cambridge, UK) containing 100-200 ml of medium. Following stirring at 40 rpm, the cell viability is estimated with trypan blue. In another format, Fibra-Cel microcarriers (Bibby Sterlin, Stone, UK) (5 g/l) is employed as follows. A cell inoculum, resuspended in 5 ml of medium, is added to the carrier (50 ml) in a 250 ml Erlenmeyer flask and left stationary, with occasional agitation, for 1 to 4 h. The medium is then replaced with 50 ml of fresh medium and shaking initiated. For virus production, cells are allowed to grow to about 80% confluence, after which time the medium is replaced (to 25% of the final volume) and adenovirus added at an MOI of 0.05. Cultures are left stationary overnight, following which the volume is increased to 100% and shaking commenced for another 72 h.

Other than the requirement that the adenovirus vector be replication defective, or at least conditionally defective, the nature of the adenovirus vector is not believed to be crucial to the successful practice of the invention. The adenovirus may be of any of the 42 different known serotypes or subgroups A-F. Adenovirus type 5 of subgroup

5 C is the preferred starting material in order to obtain a conditional replication-defective adenovirus vector for use in the present invention, since Adenovirus type 5 is a human adenovirus about which a great deal of biochemical and genetic information is known, and it has historically been used for most constructions employing adenovirus as a vector.

As stated above, the typical vector according to the present invention is

10 replication defective and will not have an adenovirus E1 region. Thus, it will be most convenient to introduce the polynucleotide encoding the gene of interest at the position from which the E1-coding sequences have been removed. However, the position of insertion of the construct within the adenovirus sequences is not critical to the invention. The polynucleotide encoding the gene of interest may also be inserted in lieu of the deleted

15 E3 region in E3 replacement vectors as described by Karlsson *et al.* (1986) or in the E4 region where a helper cell line or helper virus complements the E4 defect.

Adenovirus is easy to grow and manipulate and exhibits broad host range *in vitro* and *in vivo*. This group of viruses can be obtained in high titers, *e.g.*, 10^9 - 10^{11} plaque-forming units per ml, and they are highly infective. The life cycle of adenovirus does not

20 require integration into the host cell genome. The foreign genes delivered by adenovirus vectors are episomal and, therefore, have low genotoxicity to host cells. No side effects have been reported in studies of vaccination with wild-type adenovirus (Couch *et al.*, 1963; Top *et al.*, 1971), demonstrating their safety and therapeutic potential as *in vivo* gene transfer vectors.

25 Adenovirus vectors have been used in eukaryotic gene expression (Levrero *et al.*, 1991; Gomez-Foix *et al.*, 1992) and vaccine development (Grunhaus and Horwitz, 1992; Graham and Prevec, 1992). Recently, animal studies suggested that recombinant adenovirus could be used for gene therapy (Stratford-Perricaudet and Perricaudet, 1991; Stratford-Perricaudet *et al.*, 1990; Rich *et al.*, 1993). Studies in administering recombinant

adenovirus to different tissues include trachea instillation (Rosenfeld *et al.*, 1991; Rosenfeld *et al.*, 1992), muscle injection (Ragot *et al.*, 1993), peripheral intravenous injections (Herz and Gerard, 1993) and stereotactic inoculation into the brain (Le Gal La Salle *et al.*, 1993).

5 2. RETROVIRUSES

The retroviruses are a group of single-stranded RNA viruses characterized by an ability to convert their RNA to double-stranded DNA in infected cells by a process of reverse-transcription (Coffin, 1990). The resulting DNA then stably integrates into cellular chromosomes as a provirus and directs synthesis of viral proteins. The integration results
 10 in the retention of the viral gene sequences in the recipient cell and its descendants. The retroviral genome contains three genes, gag, pol, and env that code for capsid proteins, polymerase enzyme, and envelope components, respectively. A sequence found upstream from the gag gene contains a signal for packaging of the genome into virions. Two long terminal repeat (LTR) sequences are present at the 5' and 3' ends of the viral genome.
 15 These contain strong promoter and enhancer sequences and are also required for integration in the host cell genome (Coffin, 1990).

In order to construct a retroviral vector, a nucleic acid encoding one or more oligonucleotide or polynucleotide sequences of interest is inserted into the viral genome in the place of certain viral sequences to produce a virus that is replication-defective. In order
 20 to produce virions, a packaging cell line containing the gag, pol, and env genes but without the LTR and packaging components is constructed (Mann *et al.*, 1983). When a recombinant plasmid containing a cDNA, together with the retroviral LTR and packaging sequences is introduced into this cell line (by calcium phosphate precipitation for example), the packaging sequence allows the RNA transcript of the recombinant plasmid to be
 25 packaged into viral particles, which are then secreted into the culture media (Nicolas and Rubenstein, 1988; Temin, 1986; Mann *et al.*, 1983). The media containing the recombinant retroviruses is then collected, optionally concentrated, and used for gene

transfer. Retroviral vectors are able to infect a broad variety of cell types. However, integration and stable expression require the division of host cells (Paskind *et al.*, 1975).

A novel approach designed to allow specific targeting of retrovirus vectors was recently developed based on the chemical modification of a retrovirus by the chemical addition of lactose residues to the viral envelope. This modification could permit the specific infection of hepatocytes *via* sialoglycoprotein receptors.

A different approach to targeting of recombinant retroviruses was designed in which biotinylated antibodies against a retroviral envelope protein and against a specific cell receptor were used. The antibodies were coupled *via* the biotin components by using streptavidin (Roux *et al.*, 1989). Using antibodies against major histocompatibility complex class I and class II antigens, they demonstrated the infection of a variety of human cells that bore those surface antigens with an ecotropic virus *in vitro* (Roux *et al.*, 1989).

3. ADENO-ASSOCIATED VIRUSES

AAV (Ridgeway, 1988; Hermonat and Muzycska, 1984) is a parovirus, discovered as a contamination of adenoviral stocks. It is a ubiquitous virus (antibodies are present in 85% of the US human population) that has not been linked to any disease. It is also classified as a dependovirus, because its replications is dependent on the presence of a helper virus, such as adenovirus. Five serotypes have been isolated, of which AAV-2 is the best characterized. AAV has a single-stranded linear DNA that is encapsidated into capsid proteins VP1, VP2 and VP3 to form an icosahedral virion of 20 to 24 nm in diameter (Muzyczka and McLaughlin, 1988).

The AAV DNA is approximately 4.7 kilobases long. It contains two open reading frames and is flanked by two ITRs (FIG. 2). There are two major genes in the AAV genome: *rep* and *cap*. The *rep* gene codes for proteins responsible for viral replications, whereas *cap* codes for capsid protein VP1-3. Each ITR forms a T-shaped hairpin structure. These terminal repeats are the only essential *cis* components of the AAV for chromosomal integration. Therefore, the AAV can be used as a vector with all viral coding sequences removed and replaced by the cassette of genes for delivery. Three viral

promoters have been identified and named p5, p19, and p40, according to their map position. Transcription from p5 and p19 results in production of rep proteins, and transcription from p40 produces the capsid proteins (Hermonat and Muzyczka, 1984).

There are several factors that prompted researchers to study the possibility of using rAAV as an expression vector. One is that the requirements for delivering a gene to integrate into the host chromosome are surprisingly few. It is necessary to have the 145-bp ITRs, which are only 6% of the AAV genome. This leaves room in the vector to assemble a 4.5-kb DNA insertion. While this carrying capacity may prevent the AAV from delivering large genes, it is amply suited for delivering the antisense constructs of the present invention.

AAV is also a good choice of delivery vehicles due to its safety. There is a relatively complicated rescue mechanism: not only wild type adenovirus but also AAV genes are required to mobilize rAAV. Likewise, AAV is not pathogenic and not associated with any disease. The removal of viral coding sequences minimizes immune reactions to viral gene expression, and therefore, rAAV does not evoke an inflammatory response.

4. OTHER VIRAL VECTORS AS EXPRESSION CONSTRUCTS

Other viral vectors may be employed as expression constructs in the present invention for the delivery of oligonucleotide or polynucleotide sequences to a host cell. Vectors derived from viruses such as vaccinia virus (Ridgeway, 1988; Coupar *et al.*, 1988), lentiviruses, polio viruses and herpes viruses may be employed. They offer several attractive features for various mammalian cells (Friedmann, 1989; Ridgeway, 1988; Coupar *et al.*, 1988; Horwich *et al.*, 1990).

With the recent recognition of defective hepatitis B viruses, new insight was gained into the structure-function relationship of different viral sequences. *In vitro* studies showed that the virus could retain the ability for helper-dependent packaging and reverse transcription despite the deletion of up to 80% of its genome (Horwich *et al.*, 1990). This suggested that large portions of the genome could be replaced with foreign genetic material. The hepatotropism and persistence (integration) were particularly attractive

properties for liver-directed gene transfer. Chang *et al.* (1991) introduced the chloramphenicol acetyltransferase (CAT) gene into duck hepatitis B virus genome in the place of the polymerase, surface, and pre-surface coding sequences. It was cotransfected with wild-type virus into an avian hepatoma cell line. Culture media containing high titers of the recombinant virus were used to infect primary duckling hepatocytes. Stable CAT gene expression was detected for at least 24 days after transfection (Chang *et al.*, 1991).

5. NON-VIRAL VECTORS

In order to effect expression of the oligonucleotide or polynucleotide sequences of the present invention, the expression construct must be delivered into a cell.

This delivery may be accomplished *in vitro*, as in laboratory procedures for transforming cells lines, or *in vivo* or *ex vivo*, as in the treatment of certain disease states. As described above, one preferred mechanism for delivery is *via* viral infection where the expression construct is encapsulated in an infectious viral particle.

Once the expression construct has been delivered into the cell the nucleic acid encoding the desired oligonucleotide or polynucleotide sequences may be positioned and expressed at different sites. In certain embodiments, the nucleic acid encoding the construct may be stably integrated into the genome of the cell. This integration may be in the specific location and orientation *via* homologous recombination (gene replacement) or it may be integrated in a random, non-specific location (gene augmentation). In yet further embodiments, the nucleic acid may be stably maintained in the cell as a separate, episomal segment of DNA. Such nucleic acid segments or "episomes" encode sequences sufficient to permit maintenance and replication independent of or in synchronization with the host cell cycle. How the expression construct is delivered to a cell and where in the cell the nucleic acid remains is dependent on the type of expression construct employed.

In certain embodiments of the invention, the expression construct comprising one or more oligonucleotide or polynucleotide sequences may simply consist of naked recombinant DNA or plasmids. Transfer of the construct may be performed by any of the methods mentioned above which physically or chemically permeabilize the cell

membrane. This is particularly applicable for transfer *in vitro* but it may be applied to *in vivo* use as well. Dubensky *et al.* (1984) successfully injected polyomavirus DNA in the form of calcium phosphate precipitates into liver and spleen of adult and newborn mice demonstrating active viral replication and acute infection. Benvenisty and Reshef (1986) also demonstrated that direct intraperitoneal injection of calcium phosphate-precipitated plasmids results in expression of the transfected genes. It is envisioned that DNA encoding a gene of interest may also be transferred in a similar manner *in vivo* and express the gene product.

Another embodiment of the invention for transferring a naked DNA expression construct into cells may involve particle bombardment. This method depends on the ability to accelerate DNA-coated microprojectiles to a high velocity allowing them to pierce cell membranes and enter cells without killing them (Klein *et al.*, 1987). Several devices for accelerating small particles have been developed. One such device relies on a high voltage discharge to generate an electrical current, which in turn provides the motive force (Yang *et al.*, 1990). The microprojectiles used have consisted of biologically inert substances such as tungsten or gold beads.

Selected organs including the liver, skin, and muscle tissue of rats and mice have been bombarded *in vivo* (Yang *et al.*, 1990; Zelenin *et al.*, 1991). This may require surgical exposure of the tissue or cells, to eliminate any intervening tissue between the gun and the target organ, *i.e.* *ex vivo* treatment. Again, DNA encoding a particular gene may be delivered *via* this method and still be incorporated by the present invention.

ANTISENSE OLIGONUCLEOTIDES

The end result of the flow of genetic information is the synthesis of protein. DNA is transcribed by polymerases into messenger RNA and translated on the ribosome to yield a folded, functional protein. Thus there are several steps along the route where protein synthesis can be inhibited. The native DNA segment coding for a polypeptide described herein, as all such mammalian DNA strands, has two strands: a sense strand and an antisense strand held together by hydrogen bonding. The messenger RNA coding for

polypeptide has the same nucleotide sequence as the sense DNA strand except that the DNA thymidine is replaced by uridine. Thus, synthetic antisense nucleotide sequences will bind to a mRNA and inhibit expression of the protein encoded by that mRNA.

The targeting of antisense oligonucleotides to mRNA is thus one mechanism to shut down protein synthesis, and, consequently, represents a powerful and targeted therapeutic approach. For example, the synthesis of polygalacturonase and the muscarine type 2 acetylcholine receptor are inhibited by antisense oligonucleotides directed to their respective mRNA sequences (U. S. Patent 5,739,119 and U. S. Patent 5,759,829, each specifically incorporated herein by reference in its entirety). Further, examples of antisense inhibition have been demonstrated with the nuclear protein cyclin, the multiple drug resistance gene (MDG1), ICAM-1, E-selectin, STK-1, striatal GABA_A receptor and human EGF (Jaskulski *et al.*, 1988; Vasanthakumar and Ahmed, 1989; Peris *et al.*, 1998; U. S. Patent 5,801,154; U. S. Patent 5,789,573; U. S. Patent 5,718,709 and U. S. Patent 5,610,288, each specifically incorporated herein by reference in its entirety). Antisense constructs have also been described that inhibit and can be used to treat a variety of abnormal cellular proliferations, *e.g.* cancer (U. S. Patent 5,747,470; U. S. Patent 5,591,317 and U. S. Patent 5,783,683, each specifically incorporated herein by reference in its entirety).

Therefore, in exemplary embodiments, the invention provides oligonucleotide sequences that comprise all, or a portion of, any sequence that is capable of specifically binding to polynucleotide sequence described herein, or a complement thereof. In one embodiment, the antisense oligonucleotides comprise DNA or derivatives thereof. In another embodiment, the oligonucleotides comprise RNA or derivatives thereof. In a third embodiment, the oligonucleotides are modified DNAs comprising a phosphorothioated modified backbone. In a fourth embodiment, the oligonucleotide sequences comprise peptide nucleic acids or derivatives thereof. In each case, preferred compositions comprise a sequence region that is complementary, and more preferably substantially-complementary, and even more preferably, completely complementary to one or more portions of polynucleotides disclosed herein.

Selection of antisense compositions specific for a given gene sequence is based upon analysis of the chosen target sequence (*i.e.* in these illustrative examples the rat and human sequences) and determination of secondary structure, T_m , binding energy, relative stability, and antisense compositions were selected based upon their relative inability to form dimers, hairpins, or other secondary structures that would reduce or prohibit specific binding to the target mRNA in a host cell.

Highly preferred target regions of the mRNA, are those which are at or near the AUG translation initiation codon, and those sequences which were substantially complementary to 5' regions of the mRNA. These secondary structure analyses and target site selection considerations were performed using v.4 of the OLIGO primer analysis software (Rychlik, 1997) and the BLASTN 2.0.5 algorithm software (Altschul *et al.*, 1997).

The use of an antisense delivery method employing a short peptide vector, termed MPG (27 residues), is also contemplated. The MPG peptide contains a hydrophobic domain derived from the fusion sequence of HIV gp41 and a hydrophilic domain from the nuclear localization sequence of SV40 T-antigen (Morris *et al.*, 1997). It has been demonstrated that several molecules of the MPG peptide coat the antisense oligonucleotides and can be delivered into cultured mammalian cells in less than 1 hour with relatively high efficiency (90%). Further, the interaction with MPG strongly increases both the stability of the oligonucleotide to nuclease and the ability to cross the plasma membrane (Morris *et al.*, 1997).

RIBOZYMES

Although proteins traditionally have been used for catalysis of nucleic acids, another class of macromolecules has emerged as useful in this endeavor. Ribozymes are RNA-protein complexes that cleave nucleic acids in a site-specific fashion. Ribozymes have specific catalytic domains that possess endonuclease activity (Kim and Cech, 1987; Gerlach *et al.*, 1987; Forster and Symons, 1987). For example, a large number of ribozymes accelerate phosphoester transfer reactions with a high degree of specificity, often cleaving only one of several phosphoesters in an oligonucleotide substrate (Cech *et*

al., 1981; Michel and Westhof, 1990; Reinhold-Hurek and Shub, 1992). This specificity has been attributed to the requirement that the substrate bind via specific base-pairing interactions to the internal guide sequence ("IGS") of the ribozyme prior to chemical reaction.

5 Ribozyme catalysis has primarily been observed as part of sequence-specific cleavage/ligation reactions involving nucleic acids (Joyce, 1989; Cech *et al.*, 1981). For example, U. S. Patent No. 5,354,855 (specifically incorporated herein by reference) reports that certain ribozymes can act as endonucleases with a sequence specificity greater than that of known ribonucleases and approaching that of the DNA restriction enzymes. Thus, 10 sequence-specific ribozyme-mediated inhibition of gene expression may be particularly suited to therapeutic applications (Scanlon *et al.*, 1991; Sarver *et al.*, 1990). Recently, it was reported that ribozymes elicited genetic changes in some cells lines to which they were applied; the altered genes included the oncogenes *H-ras*, *c-fos* and genes of HIV. Most of this work involved the modification of a target mRNA, based on a specific mutant codon 15 that is cleaved by a specific ribozyme.

Six basic varieties of naturally-occurring enzymatic RNAs are known presently. Each can catalyze the hydrolysis of RNA phosphodiester bonds *in trans* (and thus can cleave other RNA molecules) under physiological conditions. In general, enzymatic nucleic acids act by first binding to a target RNA. Such binding occurs through 20 the target binding portion of a enzymatic nucleic acid which is held in close proximity to an enzymatic portion of the molecule that acts to cleave the target RNA. Thus, the enzymatic nucleic acid first recognizes and then binds a target RNA through complementary base-pairing, and once bound to the correct site, acts enzymatically to cut the target RNA. Strategic cleavage of such a target RNA will destroy its ability to direct synthesis of an 25 encoded protein. After an enzymatic nucleic acid has bound and cleaved its RNA target, it is released from that RNA to search for another target and can repeatedly bind and cleave new targets.

The enzymatic nature of a ribozyme is advantageous over many technologies, such as antisense technology (where a nucleic acid molecule simply binds to

a nucleic acid target to block its translation) since the concentration of ribozyme necessary to affect a therapeutic treatment is lower than that of an antisense oligonucleotide. This advantage reflects the ability of the ribozyme to act enzymatically. Thus, a single ribozyme molecule is able to cleave many molecules of target RNA. In addition, the
 5 ribozyme is a highly specific inhibitor, with the specificity of inhibition depending not only on the base pairing mechanism of binding to the target RNA, but also on the mechanism of target RNA cleavage. Single mismatches, or base-substitutions, near the site of cleavage can completely eliminate catalytic activity of a ribozyme. Similar mismatches in antisense molecules do not prevent their action (Woolf *et al.*, 1992). Thus, the specificity of action
 10 of a ribozyme is greater than that of an antisense oligonucleotide binding the same RNA site.

The enzymatic nucleic acid molecule may be formed in a hammerhead, hairpin, a hepatitis δ virus, group I intron or RNaseP RNA (in association with an RNA guide sequence) or Neurospora VS RNA motif. Examples of hammerhead motifs are
 15 described by Rossi *et al.* (1992). Examples of hairpin motifs are described by Hampel *et al.* (Eur. Pat. Appl. Publ. No. EP 0360257), Hampel and Tritz (1989), Hampel *et al.* (1990) and U. S. Patent 5,631,359 (specifically incorporated herein by reference). An example of the hepatitis δ virus motif is described by Perrotta and Been (1992); an example of the RNaseP motif is described by Guerrier-Takada *et al.* (1983); Neurospora VS RNA
 20 ribozyme motif is described by Collins (Saville and Collins, 1990; Saville and Collins, 1991; Collins and Olive, 1993); and an example of the Group I intron is described in (U. S. Patent 4,987,071, specifically incorporated herein by reference). All that is important in an enzymatic nucleic acid molecule of this invention is that it has a specific substrate binding site which is complementary to one or more of the target gene RNA regions, and that it
 25 have nucleotide sequences within or surrounding that substrate binding site which impart an RNA cleaving activity to the molecule. Thus the ribozyme constructs need not be limited to specific motifs mentioned herein.

In certain embodiments, it may be important to produce enzymatic cleaving agents which exhibit a high degree of specificity for the RNA of a desired target, such as

one of the sequences disclosed herein. The enzymatic nucleic acid molecule is preferably targeted to a highly conserved sequence region of a target mRNA. Such enzymatic nucleic acid molecules can be delivered exogenously to specific cells as required. Alternatively, the ribozymes can be expressed from DNA or RNA vectors that are delivered to specific
5 cells.

Small enzymatic nucleic acid motifs (*e.g.*, of the hammerhead or the hairpin structure) may also be used for exogenous delivery. The simple structure of these molecules increases the ability of the enzymatic nucleic acid to invade targeted regions of the mRNA structure. Alternatively, catalytic RNA molecules can be expressed within cells
10 from eukaryotic promoters (*e.g.*, Scanlon *et al.*, 1991; Kashani-Sabet *et al.*, 1992; Dropulic *et al.*, 1992; Weerasinghe *et al.*, 1991; Ojwang *et al.*, 1992; Chen *et al.*, 1992; Sarver *et al.*, 1990). Those skilled in the art realize that any ribozyme can be expressed in eukaryotic cells from the appropriate DNA vector. The activity of such ribozymes can be augmented by their release from the primary transcript by a second ribozyme (Int. Pat. Appl. Publ. No.
15 WO 93/23569, and Int. Pat. Appl. Publ. No. WO 94/02595, both hereby incorporated by reference; Ohkawa *et al.*, 1992; Taira *et al.*, 1991; and Ventura *et al.*, 1993).

Ribozymes may be added directly, or can be complexed with cationic lipids, lipid complexes, packaged within liposomes, or otherwise delivered to target cells. The RNA or RNA complexes can be locally administered to relevant tissues *ex vivo*, or *in vivo*
20 through injection, aerosol inhalation, infusion pump or stent, with or without their incorporation in biopolymers.

Ribozymes may be designed as described in Int. Pat. Appl. Publ. No. WO 93/23569 and Int. Pat. Appl. Publ. No. WO 94/02595, each specifically incorporated herein by reference) and synthesized to be tested *in vitro* and *in vivo*, as described. Such
25 ribozymes can also be optimized for delivery. While specific examples are provided, those in the art will recognize that equivalent RNA targets in other species can be utilized when necessary.

Hammerhead or hairpin ribozymes may be individually analyzed by computer folding (Jaeger *et al.*, 1989) to assess whether the ribozyme sequences fold into

the appropriate secondary structure. Those ribozymes with unfavorable intramolecular interactions between the binding arms and the catalytic core are eliminated from consideration. Varying binding arm lengths can be chosen to optimize activity. Generally, at least 5 or so bases on each arm are able to bind to, or otherwise interact with, the target

5 RNA.

Ribozymes of the hammerhead or hairpin motif may be designed to anneal to various sites in the mRNA message, and can be chemically synthesized. The method of synthesis used follows the procedure for normal RNA synthesis as described in Usman *et al.* (1987) and in Scaringe *et al.* (1990) and makes use of common nucleic acid

10 protecting and coupling groups, such as dimethoxytrityl at the 5'-end, and phosphoramidites at the 3'-end. Average stepwise coupling yields are typically >98%. Hairpin ribozymes may be synthesized in two parts and annealed to reconstruct an active ribozyme (Chowrira and Burke, 1992). Ribozymes may be modified extensively to enhance stability by modification with nuclease resistant groups, for example, 2'-amino, 2'-

15 C-allyl, 2'-fluoro, 2'-o-methyl, 2'-H (for a review see *e.g.*, Usman and Cedergren, 1992). Ribozymes may be purified by gel electrophoresis using general methods or by high pressure liquid chromatography and resuspended in water.

Ribozyme activity can be optimized by altering the length of the ribozyme binding arms, or chemically synthesizing ribozymes with modifications that prevent their

20 degradation by serum ribonucleases (see *e.g.*, Int. Pat. Appl. Publ. No. WO 92/07065; Perrault *et al.*, 1990; Pieken *et al.*, 1991; Usman and Cedergren, 1992; Int. Pat. Appl. Publ. No. WO 93/15187; Int. Pat. Appl. Publ. No. WO 91/03162; Eur. Pat. Appl. Publ. No. 92110298.4; U. S. Patent 5,334,711; and Int. Pat. Appl. Publ. No. WO 94/13688, which describe various chemical modifications that can be made to the sugar moieties of

25 enzymatic RNA molecules), modifications which enhance their efficacy in cells, and removal of stem II bases to shorten RNA synthesis times and reduce chemical requirements.

Sullivan *et al.* (Int. Pat. Appl. Publ. No. WO 94/02595) describes the general methods for delivery of enzymatic RNA molecules. Ribozymes may be

administered to cells by a variety of methods known to those familiar to the art, including, but not restricted to, encapsulation in liposomes, by iontophoresis, or by incorporation into other vehicles, such as hydrogels, cyclodextrins, biodegradable nanocapsules, and bioadhesive microspheres. For some indications, ribozymes may be directly delivered *ex vivo* to cells or tissues with or without the aforementioned vehicles. Alternatively, the RNA/vehicle combination may be locally delivered by direct inhalation, by direct injection or by use of a catheter, infusion pump or stent. Other routes of delivery include, but are not limited to, intravascular, intramuscular, subcutaneous or joint injection, aerosol inhalation, oral (tablet or pill form), topical, systemic, ocular, intraperitoneal and/or intrathecal delivery. More detailed descriptions of ribozyme delivery and administration are provided in Int. Pat. Appl. Publ. No. WO 94/02595 and Int. Pat. Appl. Publ. No. WO 93/23569, each specifically incorporated herein by reference.

Another means of accumulating high concentrations of a ribozyme(s) within cells is to incorporate the ribozyme-encoding sequences into a DNA expression vector.

Transcription of the ribozyme sequences are driven from a promoter for eukaryotic RNA polymerase I (pol I), RNA polymerase II (pol II), or RNA polymerase III (pol III). Transcripts from pol II or pol III promoters will be expressed at high levels in all cells; the levels of a given pol II promoter in a given cell type will depend on the nature of the gene regulatory sequences (enhancers, silencers, *etc.*) present nearby. Prokaryotic RNA polymerase promoters may also be used, providing that the prokaryotic RNA polymerase enzyme is expressed in the appropriate cells (Elroy-Stein and Moss, 1990; Gao and Huang, 1993; Lieber *et al.*, 1993; Zhou *et al.*, 1990). Ribozymes expressed from such promoters can function in mammalian cells (*e.g.* Kashani-Saber *et al.*, 1992; Ojwang *et al.*, 1992; Chen *et al.*, 1992; Yu *et al.*, 1993; L'Huillier *et al.*, 1992; Lisiewicz *et al.*, 1993). Such transcription units can be incorporated into a variety of vectors for introduction into mammalian cells, including but not restricted to, plasmid DNA vectors, viral DNA vectors (such as adenovirus or adeno-associated vectors), or viral RNA vectors (such as retroviral, semliki forest virus, sindbis virus vectors).

Ribozymes may be used as diagnostic tools to examine genetic drift and mutations within diseased cells. They can also be used to assess levels of the target RNA molecule. The close relationship between ribozyme activity and the structure of the target RNA allows the detection of mutations in any region of the molecule which alters the base-pairing and three-dimensional structure of the target RNA. By using multiple ribozymes, one may map nucleotide changes which are important to RNA structure and function *in vitro*, as well as in cells and tissues. Cleavage of target RNAs with ribozymes may be used to inhibit gene expression and define the role (essentially) of specified gene products in the progression of disease. In this manner, other genetic targets may be defined as important mediators of the disease. These studies will lead to better treatment of the disease progression by affording the possibility of combinational therapies (*e.g.*, multiple ribozymes targeted to different genes, ribozymes coupled with known small molecule inhibitors, or intermittent treatment with combinations of ribozymes and/or other chemical or biological molecules). Other *in vitro* uses of ribozymes are well known in the art, and include detection of the presence of mRNA associated with an IL-5 related condition. Such RNA is detected by determining the presence of a cleavage product after treatment with a ribozyme using standard methodology.

PEPTIDE NUCLEIC ACIDS

In certain embodiments, the inventors contemplate the use of peptide nucleic acids (PNAs) in the practice of the methods of the invention. PNA is a DNA mimic in which the nucleobases are attached to a pseudopeptide backbone (Good and Nielsen, 1997). PNA is able to be utilized in a number methods that traditionally have used RNA or DNA. Often PNA sequences perform better in techniques than the corresponding RNA or DNA sequences and have utilities that are not inherent to RNA or DNA. A review of PNA including methods of making, characteristics of, and methods of using, is provided by Corey (1997) and is incorporated herein by reference. As such, in certain embodiments, one may prepare PNA sequences that are complementary to one or more portions of the ACE mRNA sequence, and such PNA compositions may be used to regulate, alter,

decrease, or reduce the translation of ACE-specific mRNA, and thereby alter the level of ACE activity in a host cell to which such PNA compositions have been administered.

PNAs have 2-aminoethyl-glycine linkages replacing the normal phosphodiester backbone of DNA (Nielsen *et al.*, 1991; Hanvey *et al.*, 1992; Hyrup and Nielsen, 1996; Neilsen, 1996). This chemistry has three important consequences: firstly, in contrast to DNA or phosphorothioate oligonucleotides, PNAs are neutral molecules; secondly, PNAs are achiral, which avoids the need to develop a stereoselective synthesis; and thirdly, PNA synthesis uses standard Boc (Dueholm *et al.*, 1994) or Fmoc (Thomson *et al.*, 1995) protocols for solid-phase peptide synthesis, although other methods, including a modified Merrifield method, have been used (Christensen *et al.*, 1995).

PNA monomers or ready-made oligomers are commercially available from PerSeptive Biosystems (Framingham, MA). PNA syntheses by either Boc or Fmoc protocols are straightforward using manual or automated protocols (Norton *et al.*, 1995). The manual protocol lends itself to the production of chemically modified PNAs or the simultaneous synthesis of families of closely related PNAs.

As with peptide synthesis, the success of a particular PNA synthesis will depend on the properties of the chosen sequence. For example, while in theory PNAs can incorporate any combination of nucleotide bases, the presence of adjacent purines can lead to deletions of one or more residues in the product. In expectation of this difficulty, it is suggested that, in producing PNAs with adjacent purines, one should repeat the coupling of residues likely to be added inefficiently. This should be followed by the purification of PNAs by reverse-phase high-pressure liquid chromatography (Norton *et al.*, 1995) providing yields and purity of product similar to those observed during the synthesis of peptides.

Modifications of PNAs for a given application may be accomplished by coupling amino acids during solid-phase synthesis or by attaching compounds that contain a carboxylic acid group to the exposed N-terminal amine. Alternatively, PNAs can be modified after synthesis by coupling to an introduced lysine or cysteine. The ease with which PNAs can be modified facilitates optimization for better solubility or for specific

functional requirements. Once synthesized, the identity of PNAs and their derivatives can be confirmed by mass spectrometry. Several studies have made and utilized modifications of PNAs (Norton *et al.*, 1995; Haaima *et al.*, 1996; Stetsenko *et al.*, 1996; Petersen *et al.*, 1995; Ulmann *et al.*, 1996; Koch *et al.*, 1995; Orum *et al.*, 1995; Footer *et al.*, 1996; Griffith *et al.*, 1995; Kremisky *et al.*, 1996; Pardridge *et al.*, 1995; Boffa *et al.*, 1995; Landsdorp *et al.*, 1996; Gambacorti-Passerini *et al.*, 1996; Armitage *et al.*, 1997; Seeger *et al.*, 1997; Ruskowski *et al.*, 1997). U.S. Patent No. 5,700,922 discusses PNA-DNA-PNA chimeric molecules and their uses in diagnostics, modulating protein in organisms, and treatment of conditions susceptible to therapeutics.

In contrast to DNA and RNA, which contain negatively charged linkages, the PNA backbone is neutral. In spite of this dramatic alteration, PNAs recognize complementary DNA and RNA by Watson-Crick pairing (Egholm *et al.*, 1993), validating the initial modeling by Nielsen *et al.* (1991). PNAs lack 3' to 5' polarity and can bind in either parallel or antiparallel fashion, with the antiparallel mode being preferred (Egholm *et al.*, 1993).

Hybridization of DNA oligonucleotides to DNA and RNA is destabilized by electrostatic repulsion between the negatively charged phosphate backbones of the complementary strands. By contrast, the absence of charge repulsion in PNA-DNA or PNA-RNA duplexes increases the melting temperature (T_m) and reduces the dependence of T_m on the concentration of mono- or divalent cations (Nielsen *et al.*, 1991). The enhanced rate and affinity of hybridization are significant because they are responsible for the surprising ability of PNAs to perform strand invasion of complementary sequences within relaxed double-stranded DNA. In addition, the efficient hybridization at inverted repeats suggests that PNAs can recognize secondary structure effectively within double-stranded DNA. Enhanced recognition also occurs with PNAs immobilized on surfaces, and Wang *et al.* have shown that support-bound PNAs can be used to detect hybridization events (Wang *et al.*, 1996).

One might expect that tight binding of PNAs to complementary sequences would also increase binding to similar (but not identical) sequences, reducing the sequence

specificity of PNA recognition. As with DNA hybridization, however, selective recognition can be achieved by balancing oligomer length and incubation temperature. Moreover, selective hybridization of PNAs is encouraged by PNA-DNA hybridization being less tolerant of base mismatches than DNA-DNA hybridization. For example, a
 5 single mismatch within a 16 bp PNA-DNA duplex can reduce the T_m by up to 15°C (Egholm *et al.*, 1993). This high level of discrimination has allowed the development of several PNA-based strategies for the analysis of point mutations (Wang *et al.*, 1996; Carlsson *et al.*, 1996; Thiede *et al.*, 1996; Webb and Hurskainen, 1996; Perry-O'Keefe *et al.*, 1996).

10 High-affinity binding provides clear advantages for molecular recognition and the development of new applications for PNAs. For example, 11-13 nucleotide PNAs inhibit the activity of telomerase, a ribonucleo-protein that extends telomere ends using an essential RNA template, while the analogous DNA oligomers do not (Norton *et al.*, 1996).

Neutral PNAs are more hydrophobic than analogous DNA oligomers, and
 15 this can lead to difficulty solubilizing them at neutral pH, especially if the PNAs have a high purine content or if they have the potential to form secondary structures. Their solubility can be enhanced by attaching one or more positive charges to the PNA termini (Nielsen *et al.*, 1991).

Findings by Allfrey and colleagues suggest that strand invasion will occur
 20 spontaneously at sequences within chromosomal DNA (Boffa *et al.*, 1995; Boffa *et al.*, 1996). These studies targeted PNAs to triplet repeats of the nucleotides CAG and used this recognition to purify transcriptionally active DNA (Boffa *et al.*, 1995) and to inhibit transcription (Boffa *et al.*, 1996). This result suggests that if PNAs can be delivered within cells then they will have the potential to be general sequence-specific regulators of gene
 25 expression. Studies and reviews concerning the use of PNAs as antisense and anti-gene agents include Nielsen *et al.* (1993b), Hanvey *et al.* (1992), and Good and Nielsen (1997). Koppelhus *et al.* (1997) have used PNAs to inhibit HIV-1 inverse transcription, showing that PNAs may be used for antiviral therapies.

Methods of characterizing the antisense binding properties of PNAs are discussed in Rose (1993) and Jensen *et al.* (1997). Rose uses capillary gel electrophoresis to determine binding of PNAs to their complementary oligonucleotide, measuring the relative binding kinetics and stoichiometry. Similar types of measurements were made by

5 Jensen *et al.* using BIAcore™ technology.

Other applications of PNAs include use in DNA strand invasion (Nielsen *et al.*, 1991), antisense inhibition (Hanvey *et al.*, 1992), mutational analysis (Orum *et al.*, 1993), enhancers of transcription (Mollegaard *et al.*, 1994), nucleic acid purification (Orum *et al.*, 1995), isolation of transcriptionally active genes (Boffa *et al.*, 1995), blocking of

10 transcription factor binding (Vickers *et al.*, 1995), genome cleavage (Veselkov *et al.*, 1996), biosensors (Wang *et al.*, 1996), *in situ* hybridization (Thisted *et al.*, 1996), and in a alternative to Southern blotting (Perry-O'Keefe, 1996).

POLYPEPTIDE COMPOSITIONS

The present invention, in other aspects, provides polypeptide compositions.

15 Generally, a polypeptide of the invention will be an isolated polypeptide (or an epitope, variant, or active fragment thereof) derived from a mammalian species. Preferably, the polypeptide is encoded by a polynucleotide sequence disclosed herein or a sequence which hybridizes under moderately stringent conditions to a polynucleotide sequence disclosed herein. Alternatively, the polypeptide may be defined as a polypeptide which comprises a

20 contiguous amino acid sequence from an amino acid sequence disclosed herein, or which polypeptide comprises an entire amino acid sequence disclosed herein.

In the present invention, a polypeptide composition is also understood to comprise one or more polypeptides that are immunologically reactive with antibodies generated against a polypeptide of the invention, particularly a polypeptide having the

25 amino acid sequence disclosed in SEQ ID NO:110, 112, 114, 152, 155, 156, 159, 161, 165, 166, 169, 170, 172, 174, 176, 225, 226-251, 252, 338-344, 346, 348 and 350, or to active fragments, or to variants or biological functional equivalents thereof.

Likewise, a polypeptide composition of the present invention is understood to comprise one or more polypeptides that are capable of eliciting antibodies that are immunologically reactive with one or more polypeptides encoded by one or more contiguous nucleic acid sequences contained in SEQ ID NO:1-109, 111, 113, 115-151, 153, 154, 157, 158, 160, 162-164, 167, 168, 171, 173, 175, 177-224, 255-337, 345, 347 and 349, or to active fragments, or to variants thereof, or to one or more nucleic acid sequences which hybridize to one or more of these sequences under conditions of moderate to high stringency. Particularly illustrative polypeptides include the amino acid sequence disclosed in SEQ ID NO:110, 112, 114, 152, 155, 156, 159, 161, 165, 166, 169, 170, 172, 174, 176, 225, 226-251, 252, 338-344, 346, 348 and 350.

As used herein, an active fragment of a polypeptide includes a whole or a portion of a polypeptide which is modified by conventional techniques, *e.g.*, mutagenesis, or by addition, deletion, or substitution, but which active fragment exhibits substantially the same structure function, antigenicity, etc., as a polypeptide as described herein.

In certain illustrative embodiments, the polypeptides of the invention will comprise at least an immunogenic portion of a lung tumor protein or a variant thereof, as described herein. As noted above, a "lung tumor protein" is a protein that is expressed by lung tumor cells. Proteins that are lung tumor proteins also react detectably within an immunoassay (such as an ELISA) with antisera from a patient with lung cancer. Polypeptides as described herein may be of any length. Additional sequences derived from the native protein and/or heterologous sequences may be present, and such sequences may (but need not) possess further immunogenic or antigenic properties.

An "immunogenic portion," as used herein is a portion of a protein that is recognized (*i.e.*, specifically bound) by a B-cell and/or T-cell surface antigen receptor. Such immunogenic portions generally comprise at least 5 amino acid residues, more preferably at least 10, and still more preferably at least 20 amino acid residues of a lung tumor protein or a variant thereof. Certain preferred immunogenic portions include peptides in which an N-terminal leader sequence and/or transmembrane domain have been deleted. Other preferred immunogenic portions may contain a small N- and/or C-terminal

deletion (*e.g.*, 1-30 amino acids, preferably 5-15 amino acids), relative to the mature protein.

Immunogenic portions may generally be identified using well known techniques, such as those summarized in Paul, *Fundamental Immunology*, 3rd ed., 243-247

5 (Raven Press, 1993) and references cited therein. Such techniques include screening polypeptides for the ability to react with antigen-specific antibodies, antisera and/or T-cell lines or clones. As used herein, antisera and antibodies are "antigen-specific" if they specifically bind to an antigen (*i.e.*, they react with the protein in an ELISA or other immunoassay, and do not react detectably with unrelated proteins). Such antisera and
10 antibodies may be prepared as described herein, and using well known techniques. An immunogenic portion of a native lung tumor protein is a portion that reacts with such antisera and/or T-cells at a level that is not substantially less than the reactivity of the full length polypeptide (*e.g.*, in an ELISA and/or T-cell reactivity assay). Such immunogenic portions may react within such assays at a level that is similar to or greater than the
15 reactivity of the full length polypeptide. Such screens may generally be performed using methods well known to those of ordinary skill in the art, such as those described in Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988. For example, a polypeptide may be immobilized on a solid support and contacted with patient sera to allow binding of antibodies within the sera to the immobilized polypeptide.
20 Unbound sera may then be removed and bound antibodies detected using, for example, ¹²⁵I-labeled Protein A.

As noted above, a composition may comprise a variant of a native lung tumor protein. A polypeptide "variant," as used herein, is a polypeptide that differs from a native lung tumor protein in one or more substitutions, deletions, additions and/or
25 insertions, such that the immunogenicity of the polypeptide is not substantially diminished. In other words, the ability of a variant to react with antigen-specific antisera may be enhanced or unchanged, relative to the native protein, or may be diminished by less than 50%, and preferably less than 20%, relative to the native protein. Such variants may generally be identified by modifying one of the above polypeptide sequences and

evaluating the reactivity of the modified polypeptide with antigen-specific antibodies or antisera as described herein. Preferred variants include those in which one or more portions, such as an N-terminal leader sequence or transmembrane domain, have been removed. Other preferred variants include variants in which a small portion (*e.g.*, 1-30 amino acids, preferably 5-15 amino acids) has been removed from the N- and/or C-terminal of the mature protein.

Polypeptide variants encompassed by the present invention include those exhibiting at least about 70%, 75%, 80%, 85%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, or 99% or more identity (determined as described above) to the polypeptides disclosed herein.

Preferably, a variant contains conservative substitutions. A "conservative substitution" is one in which an amino acid is substituted for another amino acid that has similar properties, such that one skilled in the art of peptide chemistry would expect the secondary structure and hydropathic nature of the polypeptide to be substantially unchanged. Amino acid substitutions may generally be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity and/or the amphipathic nature of the residues. For example, negatively charged amino acids include aspartic acid and glutamic acid; positively charged amino acids include lysine and arginine; and amino acids with uncharged polar head groups having similar hydrophilicity values include leucine, isoleucine and valine; glycine and alanine; asparagine and glutamine; and serine, threonine, phenylalanine and tyrosine. Other groups of amino acids that may represent conservative changes include: (1) ala, pro, gly, glu, asp, gln, asn, ser, thr; (2) cys, ser, tyr, thr; (3) val, ile, leu, met, ala, phe; (4) lys, arg, his; and (5) phe, tyr, trp, his. A variant may also, or alternatively, contain nonconservative changes. In a preferred embodiment, variant polypeptides differ from a native sequence by substitution, deletion or addition of five amino acids or fewer. Variants may also (or alternatively) be modified by, for example, the deletion or addition of amino acids that have minimal influence on the immunogenicity, secondary structure and hydropathic nature of the polypeptide.

As noted above, polypeptides may comprise a signal (or leader) sequence at the N-terminal end of the protein, which co-translationally or post-translationally directs transfer of the protein. The polypeptide may also be conjugated to a linker or other sequence for ease of synthesis, purification or identification of the polypeptide (*e.g.*, poly-His), or to enhance binding of the polypeptide to a solid support. For example, a polypeptide may be conjugated to an immunoglobulin Fc region.

Polypeptides may be prepared using any of a variety of well known techniques. Recombinant polypeptides encoded by DNA sequences as described above may be readily prepared from the DNA sequences using any of a variety of expression vectors known to those of ordinary skill in the art. Expression may be achieved in any appropriate host cell that has been transformed or transfected with an expression vector containing a DNA molecule that encodes a recombinant polypeptide. Suitable host cells include prokaryotes, yeast, and higher eukaryotic cells, such as mammalian cells and plant cells. Preferably, the host cells employed are *E. coli*, yeast or a mammalian cell line such as COS or CHO. Supernatants from suitable host/vector systems which secrete recombinant protein or polypeptide into culture media may be first concentrated using a commercially available filter. Following concentration, the concentrate may be applied to a suitable purification matrix such as an affinity matrix or an ion exchange resin. Finally, one or more reverse phase HPLC steps can be employed to further purify a recombinant polypeptide.

Portions and other variants having less than about 100 amino acids, and generally less than about 50 amino acids, may also be generated by synthetic means, using techniques well known to those of ordinary skill in the art. For example, such polypeptides may be synthesized using any of the commercially available solid-phase techniques, such as the Merrifield solid-phase synthesis method, where amino acids are sequentially added to a growing amino acid chain. *See Merrifield, J. Am. Chem. Soc. 85:2149-2146, 1963.* Equipment for automated synthesis of polypeptides is commercially available from suppliers such as Perkin Elmer/Applied BioSystems Division (Foster City, CA), and may be operated according to the manufacturer's instructions.

Within certain specific embodiments, a polypeptide may be a fusion protein that comprises multiple polypeptides as described herein, or that comprises at least one polypeptide as described herein and an unrelated sequence, such as a known tumor protein. A fusion partner may, for example, assist in providing T helper epitopes (an immunological fusion partner), preferably T helper epitopes recognized by humans, or may assist in expressing the protein (an expression enhancer) at higher yields than the native recombinant protein. Certain preferred fusion partners are both immunological and expression enhancing fusion partners. Other fusion partners may be selected so as to increase the solubility of the protein or to enable the protein to be targeted to desired intracellular compartments. Still further fusion partners include affinity tags, which facilitate purification of the protein.

Fusion proteins may generally be prepared using standard techniques, including chemical conjugation. Preferably, a fusion protein is expressed as a recombinant protein, allowing the production of increased levels, relative to a non-fused protein, in an expression system. Briefly, DNA sequences encoding the polypeptide components may be assembled separately, and ligated into an appropriate expression vector. The 3' end of the DNA sequence encoding one polypeptide component is ligated, with or without a peptide linker, to the 5' end of a DNA sequence encoding the second polypeptide component so that the reading frames of the sequences are in phase. This permits translation into a single fusion protein that retains the biological activity of both component polypeptides.

A peptide linker sequence may be employed to separate the first and second polypeptide components by a distance sufficient to ensure that each polypeptide folds into its secondary and tertiary structures. Such a peptide linker sequence is incorporated into the fusion protein using standard techniques well known in the art. Suitable peptide linker sequences may be chosen based on the following factors: (1) their ability to adopt a flexible extended conformation; (2) their inability to adopt a secondary structure that could interact with functional epitopes on the first and second polypeptides; and (3) the lack of hydrophobic or charged residues that might react with the polypeptide functional epitopes. Preferred peptide linker sequences contain Gly, Asn and Ser residues. Other near neutral

amino acids, such as Thr and Ala may also be used in the linker sequence. Amino acid sequences which may be usefully employed as linkers include those disclosed in Maratea *et al.*, *Gene* 40:39-46, 1985; Murphy *et al.*, *Proc. Natl. Acad. Sci. USA* 83:8258-8262, 1986; U.S. Patent No. 4,935,233 and U.S. Patent No. 4,751,180. The linker sequence may
 5 generally be from 1 to about 50 amino acids in length. Linker sequences are not required when the first and second polypeptides have non-essential N-terminal amino acid regions that can be used to separate the functional domains and prevent steric interference.

The ligated DNA sequences are operably linked to suitable transcriptional or translational regulatory elements. The regulatory elements responsible for expression of
 10 DNA are located only 5' to the DNA sequence encoding the first polypeptides. Similarly, stop codons required to end translation and transcription termination signals are only present 3' to the DNA sequence encoding the second polypeptide.

Fusion proteins are also provided. Such proteins comprise a polypeptide as described herein together with an unrelated immunogenic protein. Preferably the
 15 immunogenic protein is capable of eliciting a recall response. Examples of such proteins include tetanus, tuberculosis and hepatitis proteins (*see*, for example, Stoute *et al.* *New Engl. J. Med.*, 336:86-91, 1997).

Within preferred embodiments, an immunological fusion partner is derived from protein D, a surface protein of the gram-negative bacterium *Haemophilus influenza B* (WO 91/18926). Preferably, a protein D derivative comprises approximately the first third
 20 of the protein (*e.g.*, the first N-terminal 100-110 amino acids), and a protein D derivative may be lipidated. Within certain preferred embodiments, the first 109 residues of a Lipoprotein D fusion partner is included on the N-terminus to provide the polypeptide with additional exogenous T-cell epitopes and to increase the expression level in *E. coli* (thus
 25 functioning as an expression enhancer). The lipid tail ensures optimal presentation of the antigen to antigen presenting cells. Other fusion partners include the non-structural protein from influenza virus, NS1 (hemagglutinin). Typically, the N-terminal 81 amino acids are used, although different fragments that include T-helper epitopes may be used.

In another embodiment, the immunological fusion partner is the protein known as LYTA, or a portion thereof (preferably a C-terminal portion). LYTA is derived from *Streptococcus pneumoniae*, which synthesizes an N-acetyl-L-alanine amidase known as amidase LYTA (encoded by the *LytA* gene; *Gene* 43:265-292, 1986). LYTA is an autolysin that specifically degrades certain bonds in the peptidoglycan backbone. The C-terminal domain of the LYTA protein is responsible for the affinity to the choline or to some choline analogues such as DEAE. This property has been exploited for the development of *E. coli* C-LYTA expressing plasmids useful for expression of fusion proteins. Purification of hybrid proteins containing the C-LYTA fragment at the amino terminus has been described (*see Biotechnology* 10:795-798, 1992). Within a preferred embodiment, a repeat portion of LYTA may be incorporated into a fusion protein. A repeat portion is found in the C-terminal region starting at residue 178. A particularly preferred repeat portion incorporates residues 188-305.

In general, polypeptides (including fusion proteins) and polynucleotides as described herein are isolated. An "isolated" polypeptide or polynucleotide is one that is removed from its original environment. For example, a naturally-occurring protein is isolated if it is separated from some or all of the coexisting materials in the natural system. Preferably, such polypeptides are at least about 90% pure, more preferably at least about 95% pure and most preferably at least about 99% pure. A polynucleotide is considered to be isolated if, for example, it is cloned into a vector that is not a part of the natural environment.

BINDING AGENTS

The present invention further provides agents, such as antibodies and antigen-binding fragments thereof, that specifically bind to a lung tumor protein. As used herein, an antibody, or antigen-binding fragment thereof, is said to "specifically bind" to a lung tumor protein if it reacts at a detectable level (within, for example, an ELISA) with a lung tumor protein, and does not react detectably with unrelated proteins under similar conditions. As used herein, "binding" refers to a noncovalent association between two

separate molecules such that a complex is formed. The ability to bind may be evaluated by, for example, determining a binding constant for the formation of the complex. The binding constant is the value obtained when the concentration of the complex is divided by the product of the component concentrations. In general, two compounds are said to

5 "bind," in the context of the present invention, when the binding constant for complex formation exceeds about 10^3 L/mol. The binding constant may be determined using methods well known in the art.

Binding agents may be further capable of differentiating between patients with and without a cancer, such as lung cancer, using the representative assays provided

10 herein. In other words, antibodies or other binding agents that bind to a lung tumor protein will generate a signal indicating the presence of a cancer in at least about 20% of patients with the disease, and will generate a negative signal indicating the absence of the disease in at least about 90% of individuals without the cancer. To determine whether a binding agent satisfies this requirement, biological samples (e.g., blood, sera, sputum, urine and/or

15 tumor biopsies) from patients with and without a cancer (as determined using standard clinical tests) may be assayed as described herein for the presence of polypeptides that bind to the binding agent. It will be apparent that a statistically significant number of samples with and without the disease should be assayed. Each binding agent should satisfy the above criteria; however, those of ordinary skill in the art will recognize that binding agents

20 may be used in combination to improve sensitivity.

Any agent that satisfies the above requirements may be a binding agent. For example, a binding agent may be a ribosome, with or without a peptide component, an RNA molecule or a polypeptide. In a preferred embodiment, a binding agent is an antibody or an antigen-binding fragment thereof. Antibodies may be prepared by any of a

25 variety of techniques known to those of ordinary skill in the art. *See, e.g.,* Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988. In general, antibodies can be produced by cell culture techniques, including the generation of monoclonal antibodies as described herein, or via transfection of antibody genes into suitable bacterial or mammalian cell hosts, in order to allow for the production of

recombinant antibodies. In one technique, an immunogen comprising the polypeptide is initially injected into any of a wide variety of mammals (*e.g.*, mice, rats, rabbits, sheep or goats). In this step, the polypeptides of this invention may serve as the immunogen without modification. Alternatively, particularly for relatively short polypeptides, a superior
 5 immune response may be elicited if the polypeptide is joined to a carrier protein, such as bovine serum albumin or keyhole limpet hemocyanin. The immunogen is injected into the animal host, preferably according to a predetermined schedule incorporating one or more booster immunizations, and the animals are bled periodically. Polyclonal antibodies specific for the polypeptide may then be purified from such antisera by, for example,
 10 affinity chromatography using the polypeptide coupled to a suitable solid support.

Monoclonal antibodies specific for an antigenic polypeptide of interest may be prepared, for example, using the technique of Kohler and Milstein, *Eur. J. Immunol.* 6:511-519, 1976, and improvements thereto. Briefly, these methods involve the preparation of immortal cell lines capable of producing antibodies having the desired
 15 specificity (*i.e.*, reactivity with the polypeptide of interest). Such cell lines may be produced, for example, from spleen cells obtained from an animal immunized as described above. The spleen cells are then immortalized by, for example, fusion with a myeloma cell fusion partner, preferably one that is syngeneic with the immunized animal. A variety of fusion techniques may be employed. For example, the spleen cells and myeloma cells may
 20 be combined with a nonionic detergent for a few minutes and then plated at low density on a selective medium that supports the growth of hybrid cells, but not myeloma cells. A preferred selection technique uses HAT (hypoxanthine, aminopterin, thymidine) selection. After a sufficient time, usually about 1 to 2 weeks, colonies of hybrids are observed. Single colonies are selected and their culture supernatants tested for binding activity
 25 against the polypeptide. Hybridomas having high reactivity and specificity are preferred.

Monoclonal antibodies may be isolated from the supernatants of growing hybridoma colonies. In addition, various techniques may be employed to enhance the yield, such as injection of the hybridoma cell line into the peritoneal cavity of a suitable vertebrate host, such as a mouse. Monoclonal antibodies may then be harvested from the

ascites fluid or the blood. Contaminants may be removed from the antibodies by conventional techniques, such as chromatography, gel filtration, precipitation, and extraction. The polypeptides of this invention may be used in the purification process in, for example, an affinity chromatography step.

5 Within certain embodiments, the use of antigen-binding fragments of antibodies may be preferred. Such fragments include Fab fragments, which may be prepared using standard techniques. Briefly, immunoglobulins may be purified from rabbit serum by affinity chromatography on Protein A bead columns (Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988) and digested by
10 papain to yield Fab and Fc fragments. The Fab and Fc fragments may be separated by affinity chromatography on protein A bead columns.

 Monoclonal antibodies of the present invention may be coupled to one or more therapeutic agents. Suitable agents in this regard include radionuclides, differentiation inducers, drugs, toxins, and derivatives thereof. Preferred radionuclides
15 include ^{90}Y , ^{123}I , ^{125}I , ^{131}I , ^{186}Re , ^{188}Re , ^{211}At , and ^{212}Bi . Preferred drugs include methotrexate, and pyrimidine and purine analogs. Preferred differentiation inducers include phorbol esters and butyric acid. Preferred toxins include ricin, abrin, diphtheria toxin, cholera toxin, gelonin, *Pseudomonas* exotoxin, Shigella toxin, and pokeweed antiviral protein.

20 A therapeutic agent may be coupled (*e.g.*, covalently bonded) to a suitable monoclonal antibody either directly or indirectly (*e.g.*, via a linker group). A direct reaction between an agent and an antibody is possible when each possesses a substituent capable of reacting with the other. For example, a nucleophilic group, such as an amino or
25 sulfhydryl group, on one may be capable of reacting with a carbonyl-containing group, such as an anhydride or an acid halide, or with an alkyl group containing a good leaving group (*e.g.*, a halide) on the other.

 Alternatively, it may be desirable to couple a therapeutic agent and an antibody via a linker group. A linker group can function as a spacer to distance an antibody from an agent in order to avoid interference with binding capabilities. A linker

group can also serve to increase the chemical reactivity of a substituent on an agent or an antibody, and thus increase the coupling efficiency. An increase in chemical reactivity may also facilitate the use of agents, or functional groups on agents, which otherwise would not be possible.

5 It will be evident to those skilled in the art that a variety of bifunctional or polyfunctional reagents, both homo- and hetero-functional (such as those described in the catalog of the Pierce Chemical Co., Rockford, IL), may be employed as the linker group. Coupling may be effected, for example, through amino groups, carboxyl groups, sulfhydryl groups or oxidized carbohydrate residues. There are numerous references describing such
10 methodology, *e.g.*, U.S. Patent No. 4,671,958, to Rodwell *et al.*

Where a therapeutic agent is more potent when free from the antibody portion of the immunoconjugates of the present invention, it may be desirable to use a linker group which is cleavable during or upon internalization into a cell. A number of different cleavable linker groups have been described. The mechanisms for the
15 intracellular release of an agent from these linker groups include cleavage by reduction of a disulfide bond (*e.g.*, U.S. Patent No. 4,489,710, to Spitler), by irradiation of a photolabile bond (*e.g.*, U.S. Patent No. 4,625,014, to Senter *et al.*), by hydrolysis of derivatized amino acid side chains (*e.g.*, U.S. Patent No. 4,638,045, to Kohn *et al.*), by serum complement-mediated hydrolysis (*e.g.*, U.S. Patent No. 4,671,958, to Rodwell *et al.*), and acid-catalyzed
20 hydrolysis (*e.g.*, U.S. Patent No. 4,569,789, to Blattler *et al.*).

It may be desirable to couple more than one agent to an antibody. In one embodiment, multiple molecules of an agent are coupled to one antibody molecule. In another embodiment, more than one type of agent may be coupled to one antibody. Regardless of the particular embodiment, immunoconjugates with more than one agent
25 may be prepared in a variety of ways. For example, more than one agent may be coupled directly to an antibody molecule, or linkers that provide multiple sites for attachment can be used. Alternatively, a carrier can be used.

A carrier may bear the agents in a variety of ways, including covalent bonding either directly or via a linker group. Suitable carriers include proteins such as

albumins (*e.g.*, U.S. Patent No. 4,507,234, to Kato *et al.*), peptides and polysaccharides such as aminodextran (*e.g.*, U.S. Patent No. 4,699,784, to Shih *et al.*). A carrier may also bear an agent by noncovalent bonding or by encapsulation, such as within a liposome vesicle (*e.g.*, U.S. Patent Nos. 4,429,008 and 4,873,088). Carriers specific for radionuclide agents include radiohalogenated small molecules and chelating compounds. For example, U.S. Patent No. 4,735,792 discloses representative radiohalogenated small molecules and their synthesis. A radionuclide chelate may be formed from chelating compounds that include those containing nitrogen and sulfur atoms as the donor atoms for binding the metal, or metal oxide, radionuclide. For example, U.S. Patent No. 4,673,562, to Davison *et al.* discloses representative chelating compounds and their synthesis.

A variety of routes of administration for the antibodies and immunoconjugates may be used. Typically, administration will be intravenous, intramuscular, subcutaneous or in the bed of a resected tumor. It will be evident that the precise dose of the antibody/immunoconjugate will vary depending upon the antibody used, the antigen density on the tumor, and the rate of clearance of the antibody.

T CELLS

Immunotherapeutic compositions may also, or alternatively, comprise T cells specific for a lung tumor protein. Such cells may generally be prepared *in vitro* or *ex vivo*, using standard procedures. For example, T cells may be isolated from bone marrow, peripheral blood, or a fraction of bone marrow or peripheral blood of a patient, using a commercially available cell separation system, such as the Isolex™ System, available from Nexell Therapeutics, Inc. (Irvine, CA; see also U.S. Patent No. 5,240,856; U.S. Patent No. 5,215,926; WO 89/06280; WO 91/16116 and WO 92/07243). Alternatively, T cells may be derived from related or unrelated humans, non-human mammals, cell lines or cultures.

T cells may be stimulated with a lung tumor polypeptide, polynucleotide encoding a lung tumor polypeptide and/or an antigen presenting cell (APC) that expresses such a polypeptide. Such stimulation is performed under conditions and for a time sufficient to permit the generation of T cells that are specific for the polypeptide.

Preferably, a lung tumor polypeptide or polynucleotide is present within a delivery vehicle, such as a microsphere, to facilitate the generation of specific T cells.

T cells are considered to be specific for a lung tumor polypeptide if the cells specifically proliferate, secrete cytokines or kill target cells coated with the polypeptide or expressing a gene encoding the polypeptide. T cell specificity may be evaluated using any of a variety of standard techniques. For example, within a chromium release assay or proliferation assay, a stimulation index of more than two fold increase in lysis and/or proliferation, compared to negative controls, indicates T cell specificity. Such assays may be performed, for example, as described in Chen *et al.*, *Cancer Res.* 54:1065-1070, 1994. Alternatively, detection of the proliferation of T cells may be accomplished by a variety of known techniques. For example, T cell proliferation can be detected by measuring an increased rate of DNA synthesis (*e.g.*, by pulse-labeling cultures of T cells with tritiated thymidine and measuring the amount of tritiated thymidine incorporated into DNA). Contact with a lung tumor polypeptide (100 ng/ml - 100 µg/ml, preferably 200 ng/ml - 25 µg/ml) for 3 - 7 days should result in at least a two fold increase in proliferation of the T cells. Contact as described above for 2-3 hours should result in activation of the T cells, as measured using standard cytokine assays in which a two fold increase in the level of cytokine release (*e.g.*, TNF or IFN-γ) is indicative of T cell activation (*see* Coligan *et al.*, *Current Protocols in Immunology*, vol. 1, Wiley Interscience (Greene 1998)). T cells that have been activated in response to a lung tumor polypeptide, polynucleotide or polypeptide-expressing APC may be CD4⁺ and/or CD8⁺. Lung tumor protein-specific T cells may be expanded using standard techniques. Within preferred embodiments, the T cells are derived from a patient, a related donor or an unrelated donor, and are administered to the patient following stimulation and expansion.

For therapeutic purposes, CD4⁺ or CD8⁺ T cells that proliferate in response to a lung tumor polypeptide, polynucleotide or APC can be expanded in number either *in vitro* or *in vivo*. Proliferation of such T cells *in vitro* may be accomplished in a variety of ways. For example, the T cells can be re-exposed to a lung tumor polypeptide, or a short peptide corresponding to an immunogenic portion of such a polypeptide, with or without

the addition of T cell growth factors, such as interleukin-2, and/or stimulator cells that synthesize a lung tumor polypeptide. Alternatively, one or more T cells that proliferate in the presence of a lung tumor protein can be expanded in number by cloning. Methods for cloning cells are well known in the art, and include limiting dilution.

5 PHARMACEUTICAL COMPOSITIONS

In additional embodiments, the present invention concerns formulation of one or more of the polynucleotide, polypeptide, T-cell and/or antibody compositions disclosed herein in pharmaceutically-acceptable solutions for administration to a cell or an animal, either alone, or in combination with one or more other modalities of therapy.

10 It will also be understood that, if desired, the nucleic acid segment, RNA, DNA or PNA compositions that express a polypeptide as disclosed herein may be administered in combination with other agents as well, such as, *e.g.*, other proteins or polypeptides or various pharmaceutically-active agents. In fact, there is virtually no limit to other components that may also be included, given that the additional agents do not cause a
15 significant adverse effect upon contact with the target cells or host tissues. The compositions may thus be delivered along with various other agents as required in the particular instance. Such compositions may be purified from host cells or other biological sources, or alternatively may be chemically synthesized as described herein. Likewise, such compositions may further comprise substituted or derivatized RNA or DNA
20 compositions.

Formulation of pharmaceutically-acceptable excipients and carrier solutions is well-known to those of skill in the art, as is the development of suitable dosing and treatment regimens for using the particular compositions described herein in a variety of treatment regimens, including *e.g.*, oral, parenteral, intravenous, intranasal, and
25 intramuscular administration and formulation.

1. ORAL DELIVERY

In certain applications, the pharmaceutical compositions disclosed herein may be delivered *via* oral administration to an animal. As such, these compositions may be formulated with an inert diluent or with an assimilable edible carrier, or they may be enclosed in hard- or soft-shell gelatin capsule, or they may be compressed into tablets, or they may be incorporated directly with the food of the diet.

The active compounds may even be incorporated with excipients and used in the form of ingestible tablets, buccal tables, troches, capsules, elixirs, suspensions, syrups, wafers, and the like (Mathiowitz *et al.*, 1997; Hwang *et al.*, 1998; U. S. Patent 5,641,515; U. S. Patent 5,580,579 and U. S. Patent 5,792,451, each specifically incorporated herein by reference in its entirety). The tablets, troches, pills, capsules and the like may also contain the following: a binder, as gum tragacanth, acacia, cornstarch, or gelatin; excipients, such as dicalcium phosphate; a disintegrating agent, such as corn starch, potato starch, alginic acid and the like; a lubricant, such as magnesium stearate; and a sweetening agent, such as sucrose, lactose or saccharin may be added or a flavoring agent, such as peppermint, oil of wintergreen, or cherry flavoring. When the dosage unit form is a capsule, it may contain, in addition to materials of the above type, a liquid carrier. Various other materials may be present as coatings or to otherwise modify the physical form of the dosage unit. For instance, tablets, pills, or capsules may be coated with shellac, sugar, or both. A syrup or elixir may contain the active compound sucrose as a sweetening agent methyl and propylparabens as preservatives, a dye and flavoring, such as cherry or orange flavor. Of course, any material used in preparing any dosage unit form should be pharmaceutically pure and substantially non-toxic in the amounts employed. In addition, the active compounds may be incorporated into sustained-release preparation and formulations.

Typically, these formulations may contain at least about 0.1% of the active compound or more, although the percentage of the active ingredient(s) may, of course, be varied and may conveniently be between about 1 or 2% and about 60% or 70% or more of the weight or volume of the total formulation. Naturally, the amount of active

compound(s) in each therapeutically useful composition may be prepared in such a way that a suitable dosage will be obtained in any given unit dose of the compound. Factors such as solubility, bioavailability, biological half-life, route of administration, product shelf life, as well as other pharmacological considerations will be contemplated by one skilled in the art of preparing such pharmaceutical formulations, and as such, a variety of dosages and treatment regimens may be desirable.

For oral administration the compositions of the present invention may alternatively be incorporated with one or more excipients in the form of a mouthwash, dentifrice, buccal tablet, oral spray, or sublingual orally-administered formulation. For example, a mouthwash may be prepared incorporating the active ingredient in the required amount in an appropriate solvent, such as a sodium borate solution (Dobell's Solution). Alternatively, the active ingredient may be incorporated into an oral solution such as one containing sodium borate, glycerin and potassium bicarbonate, or dispersed in a dentifrice, or added in a therapeutically-effective amount to a composition that may include water, binders, abrasives, flavoring agents, foaming agents, and humectants. Alternatively the compositions may be fashioned into a tablet or solution form that may be placed under the tongue or otherwise dissolved in the mouth.

2. INJECTABLE DELIVERY

In certain circumstances it will be desirable to deliver the pharmaceutical compositions disclosed herein parenterally, intravenously, intramuscularly, or even intraperitoneally as described in U. S. Patent 5,543,158; U. S. Patent 5,641,515 and U. S. Patent 5,399,363 (each specifically incorporated herein by reference in its entirety). Solutions of the active compounds as free base or pharmacologically acceptable salts may be prepared in water suitably mixed with a surfactant, such as hydroxypropylcellulose. Dispersions may also be prepared in glycerol, liquid polyethylene glycols, and mixtures thereof and in oils. Under ordinary conditions of storage and use, these preparations contain a preservative to prevent the growth of microorganisms.

The pharmaceutical forms suitable for injectable use include sterile aqueous solutions or dispersions and sterile powders for the extemporaneous preparation of sterile injectable solutions or dispersions (U. S. Patent 5,466,468, specifically incorporated herein by reference in its entirety). In all cases the form must be sterile and must be fluid to the extent that easy syringability exists. It must be stable under the conditions of manufacture and storage and must be preserved against the contaminating action of microorganisms, such as bacteria and fungi. The carrier can be a solvent or dispersion medium containing, for example, water, ethanol, polyol (*e.g.*, glycerol, propylene glycol, and liquid polyethylene glycol, and the like), suitable mixtures thereof, and/or vegetable oils. Proper fluidity may be maintained, for example, by the use of a coating, such as lecithin, by the maintenance of the required particle size in the case of dispersion and by the use of surfactants. The prevention of the action of microorganisms can be facilitated by various antibacterial and antifungal agents, for example, parabens, chlorobutanol, phenol, sorbic acid, thimerosal, and the like. In many cases, it will be preferable to include isotonic agents, for example, sugars or sodium chloride. Prolonged absorption of the injectable compositions can be brought about by the use in the compositions of agents delaying absorption, for example, aluminum monostearate and gelatin.

For parenteral administration in an aqueous solution, for example, the solution should be suitably buffered if necessary and the liquid diluent first rendered isotonic with sufficient saline or glucose. These particular aqueous solutions are especially suitable for intravenous, intramuscular, subcutaneous and intraperitoneal administration. In this connection, a sterile aqueous medium that can be employed will be known to those of skill in the art in light of the present disclosure. For example, one dosage may be dissolved in 1 ml of isotonic NaCl solution and either added to 1000 ml of hypodermoclysis fluid or injected at the proposed site of infusion, (see for example, "Remington's Pharmaceutical Sciences" 15th Edition, pages 1035-1038 and 1570-1580). Some variation in dosage will necessarily occur depending on the condition of the subject being treated. The person responsible for administration will, in any event, determine the appropriate dose for the individual subject. Moreover, for human administration, preparations should meet sterility,

pyrogenicity, and the general safety and purity standards as required by FDA Office of Biologics standards.

5 Sterile injectable solutions are prepared by incorporating the active compounds in the required amount in the appropriate solvent with various of the other ingredients enumerated above, as required, followed by filtered sterilization. Generally, dispersions are prepared by incorporating the various sterilized active ingredients into a sterile vehicle which contains the basic dispersion medium and the required other ingredients from those enumerated above. In the case of sterile powders for the preparation of sterile injectable solutions, the preferred methods of preparation are vacuum-drying and
10 freeze-drying techniques which yield a powder of the active ingredient plus any additional desired ingredient from a previously sterile-filtered solution thereof.

The compositions disclosed herein may be formulated in a neutral or salt form. Pharmaceutically-acceptable salts, include the acid addition salts (formed with the free amino groups of the protein) and which are formed with inorganic acids such as, for example, hydrochloric or phosphoric acids, or such organic acids as acetic, oxalic, tartaric,
15 mandelic, and the like. Salts formed with the free carboxyl groups can also be derived from inorganic bases such as, for example, sodium, potassium, ammonium, calcium, or ferric hydroxides, and such organic bases as isopropylamine, trimethylamine, histidine, procaine and the like. Upon formulation, solutions will be administered in a manner
20 compatible with the dosage formulation and in such amount as is therapeutically effective. The formulations are easily administered in a variety of dosage forms such as injectable solutions, drug-release capsules, and the like.

As used herein, "carrier" includes any and all solvents, dispersion media, vehicles, coatings, diluents, antibacterial and antifungal agents, isotonic and absorption
25 delaying agents, buffers, carrier solutions, suspensions, colloids, and the like. The use of such media and agents for pharmaceutical active substances is well known in the art. Except insofar as any conventional media or agent is incompatible with the active ingredient, its use in the therapeutic compositions is contemplated. Supplementary active ingredients can also be incorporated into the compositions.

The phrase "pharmaceutically-acceptable" refers to molecular entities and compositions that do not produce an allergic or similar untoward reaction when administered to a human. The preparation of an aqueous composition that contains a protein as an active ingredient is well understood in the art. Typically, such compositions
 5 are prepared as injectables, either as liquid solutions or suspensions; solid forms suitable for solution in, or suspension in, liquid prior to injection can also be prepared. The preparation can also be emulsified.

3. NASAL DELIVERY

In certain embodiments, the pharmaceutical compositions may be delivered
 10 by intranasal sprays, inhalation, and/or other aerosol delivery vehicles. Methods for delivering genes, nucleic acids, and peptide compositions directly to the lungs *via* nasal aerosol sprays has been described *e.g.*, in U. S. Patent 5,756,353 and U. S. Patent 5,804,212 (each specifically incorporated herein by reference in its entirety). Likewise, the delivery of drugs using intranasal microparticle resins (Takenaga *et al.*, 1998) and lysophosphatidyl-
 15 glycerol compounds (U. S. Patent 5,725,871, specifically incorporated herein by reference in its entirety) are also well-known in the pharmaceutical arts. Likewise, transmucosal drug delivery in the form of a polytetrafluoroethylene support matrix is described in U. S. Patent 5,780,045 (specifically incorporated herein by reference in its entirety).

4. LIPOSOME-, NANOCAPSULE-, AND MICROPARTICLE-MEDIATED DELIVERY

In certain embodiments, the inventors contemplate the use of liposomes, nanocapsules, microparticles, microspheres, lipid particles, vesicles, and the like, for the introduction of the compositions of the present invention into suitable host cells. In particular, the compositions of the present invention may be formulated for delivery either encapsulated in a lipid particle, a liposome, a vesicle, a nanosphere, or a nanoparticle or the
 25 like.

Such formulations may be preferred for the introduction of pharmaceutically-acceptable formulations of the nucleic acids or constructs disclosed

herein. The formation and use of liposomes is generally known to those of skill in the art (see for example, Couvreur *et al.*, 1977; Couvreur, 1988; Lasic, 1998; which describes the use of liposomes and nanocapsules in the targeted antibiotic therapy for intracellular bacterial infections and diseases). Recently, liposomes were developed with improved
 5 serum stability and circulation half-times (Gabizon and Papahadjopoulos, 1988; Allen and Choun, 1987; U. S. Patent 5,741,516, specifically incorporated herein by reference in its entirety). Further, various methods of liposome and liposome like preparations as potential drug carriers have been reviewed (Takakura, 1998; Chandran *et al.*, 1997; Margalit, 1995; U. S. Patent 5,567,434; U. S. Patent 5,552,157; U. S. Patent 5,565,213; U. S. Patent
 10 5,738,868 and U. S. Patent 5,795,587, each specifically incorporated herein by reference in its entirety).

Liposomes have been used successfully with a number of cell types that are normally resistant to transfection by other procedures including T cell suspensions, primary hepatocyte cultures and PC 12 cells (Renneisen *et al.*, 1990; Muller *et al.*, 1990). In
 15 addition, liposomes are free of the DNA length constraints that are typical of viral-based delivery systems. Liposomes have been used effectively to introduce genes, drugs (Heath and Martin, 1986; Heath *et al.*, 1986; Balazsovits *et al.*, 1989; Fresta and Puglisi, 1996), radiotherapeutic agents (Pikul *et al.*, 1987), enzymes (Imaizumi *et al.*, 1990a; Imaizumi
et al., 1990b), viruses (Faller and Baltimore, 1984), transcription factors and allosteric
 20 effectors (Nicolau and Gersonde, 1979) into a variety of cultured cell lines and animals. In addition, several successful clinical trials examining the effectiveness of liposome-mediated drug delivery have been completed (Lopez-Berestein *et al.*, 1985a; 1985b; Coune, 1988; Sculier *et al.*, 1988). Furthermore, several studies suggest that the use of liposomes is not associated with autoimmune responses, toxicity or gonadal localization
 25 after systemic delivery (Mori and Fukatsu, 1992).

Liposomes are formed from phospholipids that are dispersed in an aqueous medium and spontaneously form multilamellar concentric bilayer vesicles (also termed multilamellar vesicles (MLVs)). MLVs generally have diameters of from 25 nm to 4 μ m.

Sonication of MLVs results in the formation of small unilamellar vesicles (SUVs) with diameters in the range of 200 to 500 Å, containing an aqueous solution in the core.

Liposomes bear resemblance to cellular membranes and are contemplated for use in connection with the present invention as carriers for the peptide compositions.

- 5 They are widely suitable as both water- and lipid-soluble substances can be entrapped, *i.e.* in the aqueous spaces and within the bilayer itself, respectively. It is possible that the drug-bearing liposomes may even be employed for site-specific delivery of active agents by selectively modifying the liposomal formulation.

- 10 In addition to the teachings of Couvreur *et al.* (1977; 1988), the following information may be utilized in generating liposomal formulations. Phospholipids can form a variety of structures other than liposomes when dispersed in water, depending on the molar ratio of lipid to water. At low ratios the liposome is the preferred structure. The physical characteristics of liposomes depend on pH, ionic strength and the presence of divalent cations. Liposomes can show low permeability to ionic and polar substances, but
- 15 at elevated temperatures undergo a phase transition which markedly alters their permeability. The phase transition involves a change from a closely packed, ordered structure, known as the gel state, to a loosely packed, less-ordered structure, known as the fluid state. This occurs at a characteristic phase-transition temperature and results in an increase in permeability to ions, sugars and drugs.

- 20 In addition to temperature, exposure to proteins can alter the permeability of liposomes. Certain soluble proteins, such as cytochrome c, bind, deform and penetrate the bilayer, thereby causing changes in permeability. Cholesterol inhibits this penetration of proteins, apparently by packing the phospholipids more tightly. It is contemplated that the most useful liposome formations for antibiotic and inhibitor delivery will contain
- 25 cholesterol.

The ability to trap solutes varies between different types of liposomes. For example, MLVs are moderately efficient at trapping solutes, but SUVs are extremely inefficient. SUVs offer the advantage of homogeneity and reproducibility in size distribution, however, and a compromise between size and trapping efficiency is offered by

large unilamellar vesicles (LUVs). These are prepared by ether evaporation and are three to four times more efficient at solute entrapment than MLVs.

In addition to liposome characteristics, an important determinant in entrapping compounds is the physicochemical properties of the compound itself. Polar compounds are trapped in the aqueous spaces and nonpolar compounds bind to the lipid bilayer of the vesicle. Polar compounds are released through permeation or when the bilayer is broken, but nonpolar compounds remain affiliated with the bilayer unless it is disrupted by temperature or exposure to lipoproteins. Both types show maximum efflux rates at the phase transition temperature.

Liposomes interact with cells *via* four different mechanisms: endocytosis by phagocytic cells of the reticuloendothelial system such as macrophages and neutrophils; adsorption to the cell surface, either by nonspecific weak hydrophobic or electrostatic forces, or by specific interactions with cell-surface components; fusion with the plasma cell membrane by insertion of the lipid bilayer of the liposome into the plasma membrane, with simultaneous release of liposomal contents into the cytoplasm; and by transfer of liposomal lipids to cellular or subcellular membranes, or vice versa, without any association of the liposome contents. It often is difficult to determine which mechanism is operative and more than one may operate at the same time.

The fate and disposition of intravenously injected liposomes depend on their physical properties, such as size, fluidity, and surface charge. They may persist in tissues for h or days, depending on their composition, and half lives in the blood range from min to several h. Larger liposomes, such as MLVs and LUVs, are taken up rapidly by phagocytic cells of the reticuloendothelial system, but physiology of the circulatory system restrains the exit of such large species at most sites. They can exit only in places where large openings or pores exist in the capillary endothelium, such as the sinusoids of the liver or spleen. Thus, these organs are the predominate site of uptake. On the other hand, SUVs show a broader tissue distribution but still are sequestered highly in the liver and spleen. In general, this *in vivo* behavior limits the potential targeting of liposomes to only those

organs and tissues accessible to their large size. These include the blood, liver, spleen, bone marrow, and lymphoid organs.

Targeting is generally not a limitation in terms of the present invention. However, should specific targeting be desired, methods are available for this to be accomplished. Antibodies may be used to bind to the liposome surface and to direct the antibody and its drug contents to specific antigenic receptors located on a particular cell-type surface. Carbohydrate determinants (glycoprotein or glycolipid cell-surface components that play a role in cell-cell recognition, interaction and adhesion) may also be used as recognition sites as they have potential in directing liposomes to particular cell types. Mostly, it is contemplated that intravenous injection of liposomal preparations would be used, but other routes of administration are also conceivable.

Alternatively, the invention provides for pharmaceutically-acceptable nanocapsule formulations of the compositions of the present invention. Nanocapsules can generally entrap compounds in a stable and reproducible way (Henry-Michelland *et al.*, 1987; Quintanar-Guerrero *et al.*, 1998; Douglas *et al.*, 1987). To avoid side effects due to intracellular polymeric overloading, such ultrafine particles (sized around 0.1 μm) should be designed using polymers able to be degraded *in vivo*. Biodegradable polyalkylcyanoacrylate nanoparticles that meet these requirements are contemplated for use in the present invention. Such particles may be easily made, as described (Couvreux *et al.*, 1980; 1988; zur Muhlen *et al.*, 1998; Zambaux *et al.* 1998; Pinto-Alphandry *et al.*, 1995 and U. S. Patent 5,145,684, specifically incorporated herein by reference in its entirety).

VACCINES

In certain preferred embodiments of the present invention, vaccines are provided. The vaccines will generally comprise one or more pharmaceutical compositions, such as those discussed above, in combination with an immunostimulant. An immunostimulant may be any substance that enhances or potentiates an immune response (antibody and/or cell-mediated) to an exogenous antigen. Examples of immunostimulants include adjuvants, biodegradable microspheres (*e.g.*, polylactic galactide) and liposomes

(into which the compound is incorporated; *see e.g.*, Fullerton, U.S. Patent No. 4,235,877). Vaccine preparation is generally described in, for example, M.F. Powell and M.J. Newman, eds., "Vaccine Design (the subunit and adjuvant approach)," Plenum Press (NY, 1995). Pharmaceutical compositions and vaccines within the scope of the present invention may
 5 also contain other compounds, which may be biologically active or inactive. For example, one or more immunogenic portions of other tumor antigens may be present, either incorporated into a fusion polypeptide or as a separate compound, within the composition or vaccine.

Illustrative vaccines may contain DNA encoding one or more of the
 10 polypeptides as described above, such that the polypeptide is generated *in situ*. As noted above, the DNA may be present within any of a variety of delivery systems known to those of ordinary skill in the art, including nucleic acid expression systems, bacteria and viral expression systems. Numerous gene delivery techniques are well known in the art, such as those described by Rolland, *Crit. Rev. Therap. Drug Carrier Systems* 15:143-198, 1998,
 15 and references cited therein. Appropriate nucleic acid expression systems contain the necessary DNA sequences for expression in the patient (such as a suitable promoter and terminating signal). Bacterial delivery systems involve the administration of a bacterium (such as *Bacillus-Calmette-Guerrin*) that expresses an immunogenic portion of the polypeptide on its cell surface or secretes such an epitope. In a preferred embodiment, the
 20 DNA may be introduced using a viral expression system (*e.g.*, vaccinia or other pox virus, retrovirus, or adenovirus), which may involve the use of a non-pathogenic (defective), replication competent virus. Suitable systems are disclosed, for example, in Fisher-Hoch *et al.*, *Proc. Natl. Acad. Sci. USA* 86:317-321, 1989; Flexner *et al.*, *Ann. N.Y. Acad. Sci.* 569:86-103, 1989; Flexner *et al.*, *Vaccine* 8:17-21, 1990; U.S. Patent Nos. 4,603,112,
 25 4,769,330, and 5,017,487; WO 89/01973; U.S. Patent No. 4,777,127; GB 2,200,651; EP 0,345,242; WO 91/02805; Berkner, *Biotechniques* 6:616-627, 1988; Rosenfeld *et al.*, *Science* 252:431-434, 1991; Kolls *et al.*, *Proc. Natl. Acad. Sci. USA* 91:215-219, 1994; Kass-Eisler *et al.*, *Proc. Natl. Acad. Sci. USA* 90:11498-11502, 1993; Guzman *et al.*, *Circulation* 88:2838-2848, 1993; and Guzman *et al.*, *Cir. Res.* 73:1202-1207, 1993.

Techniques for incorporating DNA into such expression systems are well known to those of ordinary skill in the art. The DNA may also be "naked," as described, for example, in Ulmer *et al.*, *Science* 259:1745-1749, 1993 and reviewed by Cohen, *Science* 259:1691-1692, 1993. The uptake of naked DNA may be increased by coating the DNA onto
5 biodegradable beads, which are efficiently transported into the cells. It will be apparent that a vaccine may comprise both a polynucleotide and a polypeptide component. Such vaccines may provide for an enhanced immune response.

It will be apparent that a vaccine may contain pharmaceutically acceptable salts of the polynucleotides and polypeptides provided herein. Such salts may be prepared
10 from pharmaceutically acceptable non-toxic bases, including organic bases (*e.g.*, salts of primary, secondary and tertiary amines and basic amino acids) and inorganic bases (*e.g.*, sodium, potassium, lithium, ammonium, calcium and magnesium salts).

While any suitable carrier known to those of ordinary skill in the art may be employed in the vaccine compositions of this invention, the type of carrier will vary
15 depending on the mode of administration. Compositions of the present invention may be formulated for any appropriate manner of administration, including for example, topical, oral, nasal, intravenous, intracranial, intraperitoneal, subcutaneous or intramuscular administration. For parenteral administration, such as subcutaneous injection, the carrier preferably comprises water, saline, alcohol, a fat, a wax or a buffer. For oral
20 administration, any of the above carriers or a solid carrier, such as mannitol, lactose, starch, magnesium stearate, sodium saccharine, talcum, cellulose, glucose, sucrose, and magnesium carbonate, may be employed. Biodegradable microspheres (*e.g.*, polylactate polyglycolate) may also be employed as carriers for the pharmaceutical compositions of this invention. Suitable biodegradable microspheres are disclosed, for example, in U.S.
25 Patent Nos. 4,897,268; 5,075,109; 5,928,647; 5,811,128; 5,820,883; 5,853,763; 5,814,344 and 5,942,252. One may also employ a carrier comprising the particulate-protein complexes described in U.S. Patent No. 5,928,647, which are capable of inducing a class I-restricted cytotoxic T lymphocyte responses in a host.

Such compositions may also comprise buffers (*e.g.*, neutral buffered saline or phosphate buffered saline), carbohydrates (*e.g.*, glucose, mannose, sucrose or dextrans), mannitol, proteins, polypeptides or amino acids such as glycine, antioxidants, bacteriostats, chelating agents such as EDTA or glutathione, adjuvants (*e.g.*, aluminum hydroxide),
 5 solutes that render the formulation isotonic, hypotonic or weakly hypertonic with the blood of a recipient, suspending agents, thickening agents and/or preservatives. Alternatively, compositions of the present invention may be formulated as a lyophilizate. Compounds may also be encapsulated within liposomes using well known technology.

Any of a variety of immunostimulants may be employed in the vaccines of
 10 this invention. For example, an adjuvant may be included. Most adjuvants contain a substance designed to protect the antigen from rapid catabolism, such as aluminum hydroxide or mineral oil, and a stimulator of immune responses, such as lipid A, *Bordetella pertussis* or *Mycobacterium tuberculosis* derived proteins. Suitable adjuvants are commercially available as, for example, Freund's Incomplete Adjuvant and Complete
 15 Adjuvant (Difco Laboratories, Detroit, MI); Merck Adjuvant 65 (Merck and Company, Inc., Rahway, NJ); AS-2 (SmithKline Beecham, Philadelphia, PA); aluminum salts such as aluminum hydroxide gel (alum) or aluminum phosphate; salts of calcium, iron or zinc; an insoluble suspension of acylated tyrosine; acylated sugars; cationically or anionically derivatized polysaccharides; polyphosphazenes; biodegradable microspheres;
 20 monophosphoryl lipid A and quil A. Cytokines, such as GM-CSF or interleukin-2, -7, or -12, may also be used as adjuvants.

Within the vaccines provided herein, the adjuvant composition is preferably designed to induce an immune response predominantly of the Th1 type. High levels of Th1-type cytokines (*e.g.*, IFN- γ , TNF α , IL-2 and IL-12) tend to favor the induction of cell
 25 mediated immune responses to an administered antigen. In contrast, high levels of Th2-type cytokines (*e.g.*, IL-4, IL-5, IL-6 and IL-10) tend to favor the induction of humoral immune responses. Following application of a vaccine as provided herein, a patient will support an immune response that includes Th1- and Th2-type responses. Within a preferred embodiment, in which a response is predominantly Th1-type, the level of Th1-

type cytokines will increase to a greater extent than the level of Th2-type cytokines. The levels of these cytokines may be readily assessed using standard assays. For a review of the families of cytokines, see Mosmann and Coffman, *Ann. Rev. Immunol.* 7:145-173, 1989.

5 Preferred adjuvants for use in eliciting a predominantly Th1-type response include, for example, a combination of monophosphoryl lipid A, preferably 3-de-O-acylated monophosphoryl lipid A (3D-MPL), together with an aluminum salt. MPL adjuvants are available from Corixa Corporation (Seattle, WA; *see* US Patent Nos. 4,436,727; 4,877,611; 4,866,034 and 4,912,094). CpG-containing oligonucleotides (in
10 which the CpG dinucleotide is unmethylated) also induce a predominantly Th1 response. Such oligonucleotides are well known and are described, for example, in WO 96/02555, WO 99/33488 and U.S. Patent Nos. 6,008,200 and 5,856,462. Immunostimulatory DNA sequences are also described, for example, by Sato *et al.*, *Science* 273:352, 1996. Another preferred adjuvant is a saponin, preferably QS21 (Aquila Biopharmaceuticals Inc.,
15 Framingham, MA), which may be used alone or in combination with other adjuvants. For example, an enhanced system involves the combination of a monophosphoryl lipid A and saponin derivative, such as the combination of QS21 and 3D-MPL as described in WO 94/00153, or a less reactogenic composition where the QS21 is quenched with cholesterol, as described in WO 96/33739. Other preferred formulations comprise an oil-in-water
20 emulsion and tocopherol. A particularly potent adjuvant formulation involving QS21, 3D-MPL and tocopherol in an oil-in-water emulsion is described in WO 95/17210.

Other preferred adjuvants include Montanide ISA 720 (Seppic, France), SAF (Chiron, California, United States), ISCOMS (CSL), MF-59 (Chiron), the SBAS series of adjuvants (*e.g.*, SBAS-2 or SBAS-4, available from SmithKline Beecham,
25 Rixensart, Belgium), Detox (Corixa, Hamilton, MT), RC-529 (Corixa, Hamilton, MT) and other aminoalkyl glucosaminide 4-phosphates (AGPs), such as those described in pending U.S. Patent Application Serial Nos. 08/853,826 and 09/074,720, the disclosures of which are incorporated herein by reference in their entireties.

Any vaccine provided herein may be prepared using well known methods that result in a combination of antigen, immune response enhancer and a suitable carrier or excipient. The compositions described herein may be administered as part of a sustained release formulation (*i.e.*, a formulation such as a capsule, sponge or gel (composed of polysaccharides, for example) that effects a slow release of compound following administration). Such formulations may generally be prepared using well known technology (*see, e.g.*, Coombes *et al.*, *Vaccine* 14:1429-1438, 1996) and administered by, for example, oral, rectal or subcutaneous implantation, or by implantation at the desired target site. Sustained-release formulations may contain a polypeptide, polynucleotide or antibody dispersed in a carrier matrix and/or contained within a reservoir surrounded by a rate controlling membrane.

Carriers for use within such formulations are biocompatible, and may also be biodegradable; preferably the formulation provides a relatively constant level of active component release. Such carriers include microparticles of poly(lactide-co-glycolide), polyacrylate, latex, starch, cellulose, dextran and the like. Other delayed-release carriers include supramolecular biovectors, which comprise a non-liquid hydrophilic core (*e.g.*, a cross-linked polysaccharide or oligosaccharide) and, optionally, an external layer comprising an amphiphilic compound, such as a phospholipid (*see e.g.*, U.S. Patent No. 5,151,254 and PCT applications WO 94/20078, WO/94/23701 and WO 96/06638). The amount of active compound contained within a sustained release formulation depends upon the site of implantation, the rate and expected duration of release and the nature of the condition to be treated or prevented.

Any of a variety of delivery vehicles may be employed within pharmaceutical compositions and vaccines to facilitate production of an antigen-specific immune response that targets tumor cells. Delivery vehicles include antigen presenting cells (APCs), such as dendritic cells, macrophages, B cells, monocytes and other cells that may be engineered to be efficient APCs. Such cells may, but need not, be genetically modified to increase the capacity for presenting the antigen, to improve activation and/or maintenance of the T cell response, to have anti-tumor effects *per se* and/or to be

immunologically compatible with the receiver (*i.e.*, matched HLA haplotype). APCs may generally be isolated from any of a variety of biological fluids and organs, including tumor and peritumoral tissues, and may be autologous, allogeneic, syngeneic or xenogeneic cells.

Certain preferred embodiments of the present invention use dendritic cells or progenitors thereof as antigen-presenting cells. Dendritic cells are highly potent APCs (Banchereau and Steinman, *Nature* 392:245-251, 1998) and have been shown to be effective as a physiological adjuvant for eliciting prophylactic or therapeutic antitumor immunity (*see* Timmerman and Levy, *Ann. Rev. Med.* 50:507-529, 1999). In general, dendritic cells may be identified based on their typical shape (stellate *in situ*, with marked cytoplasmic processes (dendrites) visible *in vitro*), their ability to take up, process and present antigens with high efficiency and their ability to activate naïve T cell responses. Dendritic cells may, of course, be engineered to express specific cell-surface receptors or ligands that are not commonly found on dendritic cells *in vivo* or *ex vivo*, and such modified dendritic cells are contemplated by the present invention. As an alternative to dendritic cells, secreted vesicles antigen-loaded dendritic cells (called exosomes) may be used within a vaccine (*see* Zitvogel *et al.*, *Nature Med.* 4:594-600, 1998).

Dendritic cells and progenitors may be obtained from peripheral blood, bone marrow, tumor-infiltrating cells, peritumoral tissues-infiltrating cells, lymph nodes, spleen, skin, umbilical cord blood or any other suitable tissue or fluid. For example, dendritic cells may be differentiated *ex vivo* by adding a combination of cytokines such as GM-CSF, IL-4, IL-13 and/or TNF α to cultures of monocytes harvested from peripheral blood. Alternatively, CD34 positive cells harvested from peripheral blood, umbilical cord blood or bone marrow may be differentiated into dendritic cells by adding to the culture medium combinations of GM-CSF, IL-3, TNF α , CD40 ligand, LPS, flt3 ligand and/or other compound(s) that induce differentiation, maturation and proliferation of dendritic cells.

Dendritic cells are conveniently categorized as "immature" and "mature" cells, which allows a simple way to discriminate between two well characterized phenotypes. However, this nomenclature should not be construed to exclude all possible intermediate stages of differentiation. Immature dendritic cells are characterized as APC

with a high capacity for antigen uptake and processing, which correlates with the high expression of Fcγ receptor and mannose receptor. The mature phenotype is typically characterized by a lower expression of these markers, but a high expression of cell surface molecules responsible for T cell activation such as class I and class II MHC, adhesion molecules (e.g., CD54 and CD11) and costimulatory molecules (e.g., CD40, CD80, CD86 and 4-1BB).

APCs may generally be transfected with a polynucleotide encoding a lung tumor protein (or portion or other variant thereof) such that the lung tumor polypeptide, or an immunogenic portion thereof, is expressed on the cell surface. Such transfection may take place *ex vivo*, and a composition or vaccine comprising such transfected cells may then be used for therapeutic purposes, as described herein. Alternatively, a gene delivery vehicle that targets a dendritic or other antigen presenting cell may be administered to a patient, resulting in transfection that occurs *in vivo*. *In vivo* and *ex vivo* transfection of dendritic cells, for example, may generally be performed using any methods known in the art, such as those described in WO 97/24447, or the gene gun approach described by Mahvi *et al.*, *Immunology and cell Biology* 75:456-460, 1997. Antigen loading of dendritic cells may be achieved by incubating dendritic cells or progenitor cells with the lung tumor polypeptide, DNA (naked or within a plasmid vector) or RNA; or with antigen-expressing recombinant bacterium or viruses (e.g., vaccinia, fowlpox, adenovirus or lentivirus vectors). Prior to loading, the polypeptide may be covalently conjugated to an immunological partner that provides T cell help (e.g., a carrier molecule). Alternatively, a dendritic cell may be pulsed with a non-conjugated immunological partner, separately or in the presence of the polypeptide.

Vaccines and pharmaceutical compositions may be presented in unit-dose or multi-dose containers, such as sealed ampoules or vials. Such containers are preferably hermetically sealed to preserve sterility of the formulation until use. In general, formulations may be stored as suspensions, solutions or emulsions in oily or aqueous vehicles. Alternatively, a vaccine or pharmaceutical composition may be stored in a

freeze-dried condition requiring only the addition of a sterile liquid carrier immediately prior to use.

CANCER THERAPY

In further aspects of the present invention, the compositions described
 5 herein may be used for immunotherapy of cancer, such as lung cancer. Within such methods, pharmaceutical compositions and vaccines are typically administered to a patient. As used herein, a “patient” refers to any warm-blooded animal, preferably a human. A patient may or may not be afflicted with cancer. Accordingly, the above pharmaceutical compositions and vaccines may be used to prevent the development of a cancer or to treat a
 10 patient afflicted with a cancer. A cancer may be diagnosed using criteria generally accepted in the art, including the presence of a malignant tumor. Pharmaceutical compositions and vaccines may be administered either prior to or following surgical removal of primary tumors and/or treatment such as administration of radiotherapy or conventional chemotherapeutic drugs. Administration may be by any suitable method,
 15 including administration by intravenous, intraperitoneal, intramuscular, subcutaneous, intranasal, intradermal, anal, vaginal, topical and oral routes.

Within certain embodiments, immunotherapy may be active immunotherapy, in which treatment relies on the *in vivo* stimulation of the endogenous host immune system to react against tumors with the administration of immune response-
 20 modifying agents (such as polypeptides and polynucleotides as provided herein).

Within other embodiments, immunotherapy may be passive immunotherapy, in which treatment involves the delivery of agents with established tumor-immune reactivity (such as effector cells or antibodies) that can directly or indirectly mediate antitumor effects and does not necessarily depend on an intact host immune
 25 system. Examples of effector cells include T cells as discussed above, T lymphocytes (such as CD8⁺ cytotoxic T lymphocytes and CD4⁺ T-helper tumor-infiltrating lymphocytes), killer cells (such as Natural Killer cells and lymphokine-activated killer cells), B cells and antigen-presenting cells (such as dendritic cells and macrophages)

expressing a polypeptide provided herein. T cell receptors and antibody receptors specific for the polypeptides recited herein may be cloned, expressed and transferred into other vectors or effector cells for adoptive immunotherapy. The polypeptides provided herein may also be used to generate antibodies or anti-idiotypic antibodies (as described above and in U.S. Patent No. 4,918,164) for passive immunotherapy.

Effector cells may generally be obtained in sufficient quantities for adoptive immunotherapy by growth *in vitro*, as described herein. Culture conditions for expanding single antigen-specific effector cells to several billion in number with retention of antigen recognition *in vivo* are well known in the art. Such *in vitro* culture conditions typically use intermittent stimulation with antigen, often in the presence of cytokines (such as IL-2) and non-dividing feeder cells. As noted above, immunoreactive polypeptides as provided herein may be used to rapidly expand antigen-specific T cell cultures in order to generate a sufficient number of cells for immunotherapy. In particular, antigen-presenting cells, such as dendritic, macrophage, monocyte, fibroblast and/or B cells, may be pulsed with immunoreactive polypeptides or transfected with one or more polynucleotides using standard techniques well known in the art. For example, antigen-presenting cells can be transfected with a polynucleotide having a promoter appropriate for increasing expression in a recombinant virus or other expression system. Cultured effector cells for use in therapy must be able to grow and distribute widely, and to survive long term *in vivo*. Studies have shown that cultured effector cells can be induced to grow *in vivo* and to survive long term in substantial numbers by repeated stimulation with antigen supplemented with IL-2 (*see, for example, Cheever et al., Immunological Reviews 157:177, 1997*).

Alternatively, a vector expressing a polypeptide recited herein may be introduced into antigen presenting cells taken from a patient and clonally propagated *ex vivo* for transplant back into the same patient. Transfected cells may be reintroduced into the patient using any means known in the art, preferably in sterile form by intravenous, intracavitary, intraperitoneal or intratumor administration.

Routes and frequency of administration of the therapeutic compositions described herein, as well as dosage, will vary from individual to individual, and may be readily established using standard techniques. In general, the pharmaceutical compositions and vaccines may be administered by injection (*e.g.*, intracutaneous, intramuscular, intravenous or subcutaneous), intranasally (*e.g.*, by aspiration) or orally. Preferably, between 1 and 10 doses may be administered over a 52 week period. Preferably, 6 doses are administered, at intervals of 1 month, and booster vaccinations may be given periodically thereafter. Alternate protocols may be appropriate for individual patients. A suitable dose is an amount of a compound that, when administered as described above, is capable of promoting an anti-tumor immune response, and is at least 10-50% above the basal (*i.e.*, untreated) level. Such response can be monitored by measuring the anti-tumor antibodies in a patient or by vaccine-dependent generation of cytolytic effector cells capable of killing the patient's tumor cells *in vitro*. Such vaccines should also be capable of causing an immune response that leads to an improved clinical outcome (*e.g.*, more frequent remissions, complete or partial or longer disease-free survival) in vaccinated patients as compared to non-vaccinated patients. In general, for pharmaceutical compositions and vaccines comprising one or more polypeptides, the amount of each polypeptide present in a dose ranges from about 25 μ g to 5 mg per kg of host. Suitable dose sizes will vary with the size of the patient, but will typically range from about 0.1 mL to about 5 mL.

In general, an appropriate dosage and treatment regimen provides the active compound(s) in an amount sufficient to provide therapeutic and/or prophylactic benefit. Such a response can be monitored by establishing an improved clinical outcome (*e.g.*, more frequent remissions, complete or partial, or longer disease-free survival) in treated patients as compared to non-treated patients. Increases in preexisting immune responses to a lung tumor protein generally correlate with an improved clinical outcome. Such immune responses may generally be evaluated using standard proliferation, cytotoxicity or cytokine assays, which may be performed using samples obtained from a patient before and after treatment.

CANCER DETECTION AND DIAGNOSIS

In general, a cancer may be detected in a patient based on the presence of one or more lung tumor proteins and/or polynucleotides encoding such proteins in a biological sample (for example, blood, sera, sputum urine and/or tumor biopsies) obtained from the patient. In other words, such proteins may be used as markers to indicate the presence or absence of a cancer such as lung cancer. In addition, such proteins may be useful for the detection of other cancers. The binding agents provided herein generally permit detection of the level of antigen that binds to the agent in the biological sample. Polynucleotide primers and probes may be used to detect the level of mRNA encoding a tumor protein, which is also indicative of the presence or absence of a cancer. In general, a lung tumor sequence should be present at a level that is at least three fold higher in tumor tissue than in normal tissue

There are a variety of assay formats known to those of ordinary skill in the art for using a binding agent to detect polypeptide markers in a sample. *See, e.g.,* Harlow and Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory, 1988. In general, the presence or absence of a cancer in a patient may be determined by (a) contacting a biological sample obtained from a patient with a binding agent; (b) detecting in the sample a level of polypeptide that binds to the binding agent; and (c) comparing the level of polypeptide with a predetermined cut-off value.

In a preferred embodiment, the assay involves the use of binding agent immobilized on a solid support to bind to and remove the polypeptide from the remainder of the sample. The bound polypeptide may then be detected using a detection reagent that contains a reporter group and specifically binds to the binding agent/polypeptide complex. Such detection reagents may comprise, for example, a binding agent that specifically binds to the polypeptide or an antibody or other agent that specifically binds to the binding agent, such as an anti-immunoglobulin, protein G, protein A or a lectin. Alternatively, a competitive assay may be utilized, in which a polypeptide is labeled with a reporter group and allowed to bind to the immobilized binding agent after incubation of the binding agent with the sample. The extent to which components of the sample inhibit the binding of the

labeled polypeptide to the binding agent is indicative of the reactivity of the sample with the immobilized binding agent. Suitable polypeptides for use within such assays include full length lung tumor proteins and portions thereof to which the binding agent binds, as described above.

5 The solid support may be any material known to those of ordinary skill in the art to which the tumor protein may be attached. For example, the solid support may be a test well in a microtiter plate or a nitrocellulose or other suitable membrane. Alternatively, the support may be a bead or disc, such as glass, fiberglass, latex or a plastic material such as polystyrene or polyvinylchloride. The support may also be a magnetic
10 particle or a fiber optic sensor, such as those disclosed, for example, in U.S. Patent No. 5,359,681. The binding agent may be immobilized on the solid support using a variety of techniques known to those of skill in the art, which are amply described in the patent and scientific literature. In the context of the present invention, the term "immobilization" refers to both noncovalent association, such as adsorption, and covalent attachment (which
15 may be a direct linkage between the agent and functional groups on the support or may be a linkage by way of a cross-linking agent). Immobilization by adsorption to a well in a microtiter plate or to a membrane is preferred. In such cases, adsorption may be achieved by contacting the binding agent, in a suitable buffer, with the solid support for a suitable amount of time. The contact time varies with temperature, but is typically between about 1
20 hour and about 1 day. In general, contacting a well of a plastic microtiter plate (such as polystyrene or polyvinylchloride) with an amount of binding agent ranging from about 10 ng to about 10 μ g, and preferably about 100 ng to about 1 μ g, is sufficient to immobilize an adequate amount of binding agent.

25 Covalent attachment of binding agent to a solid support may generally be achieved by first reacting the support with a bifunctional reagent that will react with both the support and a functional group, such as a hydroxyl or amino group, on the binding agent. For example, the binding agent may be covalently attached to supports having an appropriate polymer coating using benzoquinone or by condensation of an aldehyde group

on the support with an amine and an active hydrogen on the binding partner (*see, e.g.,* Pierce Immunotechnology Catalog and Handbook, 1991, at A12-A13).

In certain embodiments, the assay is a two-antibody sandwich assay. This assay may be performed by first contacting an antibody that has been immobilized on a solid support, commonly the well of a microtiter plate, with the sample, such that polypeptides within the sample are allowed to bind to the immobilized antibody. Unbound sample is then removed from the immobilized polypeptide-antibody complexes and a detection reagent (preferably a second antibody capable of binding to a different site on the polypeptide) containing a reporter group is added. The amount of detection reagent that remains bound to the solid support is then determined using a method appropriate for the specific reporter group.

More specifically, once the antibody is immobilized on the support as described above, the remaining protein binding sites on the support are typically blocked. Any suitable blocking agent known to those of ordinary skill in the art, such as bovine serum albumin or Tween 20™ (Sigma Chemical Co., St. Louis, MO). The immobilized antibody is then incubated with the sample, and polypeptide is allowed to bind to the antibody. The sample may be diluted with a suitable diluent, such as phosphate-buffered saline (PBS) prior to incubation. In general, an appropriate contact time (*i.e.,* incubation time) is a period of time that is sufficient to detect the presence of polypeptide within a sample obtained from an individual with lung cancer. Preferably, the contact time is sufficient to achieve a level of binding that is at least about 95% of that achieved at equilibrium between bound and unbound polypeptide. Those of ordinary skill in the art will recognize that the time necessary to achieve equilibrium may be readily determined by assaying the level of binding that occurs over a period of time. At room temperature, an incubation time of about 30 minutes is generally sufficient.

Unbound sample may then be removed by washing the solid support with an appropriate buffer, such as PBS containing 0.1% Tween 20™. The second antibody, which contains a reporter group, may then be added to the solid support. Preferred reporter groups include those groups recited above.

The detection reagent is then incubated with the immobilized antibody-polypeptide complex for an amount of time sufficient to detect the bound polypeptide. An appropriate amount of time may generally be determined by assaying the level of binding that occurs over a period of time. Unbound detection reagent is then removed and bound

5 detection reagent is detected using the reporter group. The method employed for detecting the reporter group depends upon the nature of the reporter group. For radioactive groups, scintillation counting or autoradiographic methods are generally appropriate. Spectroscopic methods may be used to detect dyes, luminescent groups and fluorescent groups. Biotin may be detected using avidin, coupled to a different reporter group

10 (commonly a radioactive or fluorescent group or an enzyme). Enzyme reporter groups may generally be detected by the addition of substrate (generally for a specific period of time), followed by spectroscopic or other analysis of the reaction products.

To determine the presence or absence of a cancer, such as lung cancer, the signal detected from the reporter group that remains bound to the solid support is generally

15 compared to a signal that corresponds to a predetermined cut-off value. In one preferred embodiment, the cut-off value for the detection of a cancer is the average mean signal obtained when the immobilized antibody is incubated with samples from patients without the cancer. In general, a sample generating a signal that is three standard deviations above the predetermined cut-off value is considered positive for the cancer. In an alternate

20 preferred embodiment, the cut-off value is determined using a Receiver Operator Curve, according to the method of Sackett *et al.*, *Clinical Epidemiology: A Basic Science for Clinical Medicine*, Little Brown and Co., 1985, p. 106-7. Briefly, in this embodiment, the cut-off value may be determined from a plot of pairs of true positive rates (*i.e.*, sensitivity) and false positive rates (100%-specificity) that correspond to each possible cut-off value

25 for the diagnostic test result. The cut-off value on the plot that is the closest to the upper left-hand corner (*i.e.*, the value that encloses the largest area) is the most accurate cut-off value, and a sample generating a signal that is higher than the cut-off value determined by this method may be considered positive. Alternatively, the cut-off value may be shifted to the left along the plot, to minimize the false positive rate, or to the right, to minimize the

false negative rate. In general, a sample generating a signal that is higher than the cut-off value determined by this method is considered positive for a cancer.

In a related embodiment, the assay is performed in a flow-through or strip test format, wherein the binding agent is immobilized on a membrane, such as nitrocellulose. In the flow-through test, polypeptides within the sample bind to the immobilized binding agent as the sample passes through the membrane. A second, labeled binding agent then binds to the binding agent-polypeptide complex as a solution containing the second binding agent flows through the membrane. The detection of bound second binding agent may then be performed as described above. In the strip test format, one end of the membrane to which binding agent is bound is immersed in a solution containing the sample. The sample migrates along the membrane through a region containing second binding agent and to the area of immobilized binding agent. Concentration of second binding agent at the area of immobilized antibody indicates the presence of a cancer. Typically, the concentration of second binding agent at that site generates a pattern, such as a line, that can be read visually. The absence of such a pattern indicates a negative result. In general, the amount of binding agent immobilized on the membrane is selected to generate a visually discernible pattern when the biological sample contains a level of polypeptide that would be sufficient to generate a positive signal in the two-antibody sandwich assay, in the format discussed above. Preferred binding agents for use in such assays are antibodies and antigen-binding fragments thereof. Preferably, the amount of antibody immobilized on the membrane ranges from about 25 ng to about 1 μ g, and more preferably from about 50 ng to about 500 ng. Such tests can typically be performed with a very small amount of biological sample.

Of course, numerous other assay protocols exist that are suitable for use with the tumor proteins or binding agents of the present invention. The above descriptions are intended to be exemplary only. For example, it will be apparent to those of ordinary skill in the art that the above protocols may be readily modified to use lung tumor polypeptides to detect antibodies that bind to such polypeptides in a biological sample.

The detection of such lung tumor protein specific antibodies may correlate with the presence of a cancer.

A cancer may also, or alternatively, be detected based on the presence of T cells that specifically react with a lung tumor protein in a biological sample. Within certain methods, a biological sample comprising CD4⁺ and/or CD8⁺ T cells isolated from a patient is incubated with a lung tumor polypeptide, a polynucleotide encoding such a polypeptide and/or an APC that expresses at least an immunogenic portion of such a polypeptide, and the presence or absence of specific activation of the T cells is detected. Suitable biological samples include, but are not limited to, isolated T cells. For example, T cells may be isolated from a patient by routine techniques (such as by Ficoll/Hypaque density gradient centrifugation of peripheral blood lymphocytes). T cells may be incubated *in vitro* for 2-9 days (typically 4 days) at 37°C with polypeptide (*e.g.*, 5 - 25 µg/ml). It may be desirable to incubate another aliquot of a T cell sample in the absence of lung tumor polypeptide to serve as a control. For CD4⁺ T cells, activation is preferably detected by evaluating proliferation of the T cells. For CD8⁺ T cells, activation is preferably detected by evaluating cytolytic activity. A level of proliferation that is at least two fold greater and/or a level of cytolytic activity that is at least 20% greater than in disease-free patients indicates the presence of a cancer in the patient.

As noted above, a cancer may also, or alternatively, be detected based on the level of mRNA encoding a lung tumor protein in a biological sample. For example, at least two oligonucleotide primers may be employed in a polymerase chain reaction (PCR) based assay to amplify a portion of a lung tumor cDNA derived from a biological sample, wherein at least one of the oligonucleotide primers is specific for (*i.e.*, hybridizes to) a polynucleotide encoding the lung tumor protein. The amplified cDNA is then separated and detected using techniques well known in the art, such as gel electrophoresis. Similarly, oligonucleotide probes that specifically hybridize to a polynucleotide encoding a lung tumor protein may be used in a hybridization assay to detect the presence of polynucleotide encoding the tumor protein in a biological sample.

To permit hybridization under assay conditions, oligonucleotide primers and probes should comprise an oligonucleotide sequence that has at least about 60%, preferably at least about 75% and more preferably at least about 90%, identity to a portion of a polynucleotide encoding a lung tumor protein that is at least 10 nucleotides, and preferably at least 20 nucleotides, in length. Preferably, oligonucleotide primers and/or probes hybridize to a polynucleotide encoding a polypeptide described herein under moderately stringent conditions, as defined above. Oligonucleotide primers and/or probes which may be usefully employed in the diagnostic methods described herein preferably are at least 10-40 nucleotides in length. In a preferred embodiment, the oligonucleotide primers comprise at least 10 contiguous nucleotides, more preferably at least 15 contiguous nucleotides, of a DNA molecule having a sequence recited in SEQ ID NO:1-109, 111, 113, 115-151, 153, 154, 157, 158, 160, 162-164, 167, 168, 171, 173, 175, 177-224, 255-337, 345, 347 and 349. Techniques for both PCR based assays and hybridization assays are well known in the art (*see, for example, Mullis et al., Cold Spring Harbor Symp. Quant. Biol., 51:263, 1987; Erlich ed., PCR Technology, Stockton Press, NY, 1989*).

One preferred assay employs RT-PCR, in which PCR is applied in conjunction with reverse transcription. Typically, RNA is extracted from a biological sample, such as biopsy tissue, and is reverse transcribed to produce cDNA molecules. PCR amplification using at least one specific primer generates a cDNA molecule, which may be separated and visualized using, for example, gel electrophoresis. Amplification may be performed on biological samples taken from a test patient and from an individual who is not afflicted with a cancer. The amplification reaction may be performed on several dilutions of cDNA spanning two orders of magnitude. A two-fold or greater increase in expression in several dilutions of the test patient sample as compared to the same dilutions of the non-cancerous sample is typically considered positive.

In another embodiment, the compositions described herein may be used as markers for the progression of cancer. In this embodiment, assays as described above for the diagnosis of a cancer may be performed over time, and the change in the level of reactive polypeptide(s) or polynucleotide(s) evaluated. For example, the assays may be

performed every 24-72 hours for a period of 6 months to 1 year, and thereafter performed as needed. In general, a cancer is progressing in those patients in whom the level of polypeptide or polynucleotide detected increases over time. In contrast, the cancer is not progressing when the level of reactive polypeptide or polynucleotide either remains
 5 constant or decreases with time.

Certain *in vivo* diagnostic assays may be performed directly on a tumor. One such assay involves contacting tumor cells with a binding agent. The bound binding agent may then be detected directly or indirectly via a reporter group. Such binding agents may also be used in histological applications. Alternatively, polynucleotide probes may be
 10 used within such applications.

As noted above, to improve sensitivity, multiple lung tumor protein markers may be assayed within a given sample. It will be apparent that binding agents specific for different proteins provided herein may be combined within a single assay. Further, multiple primers or probes may be used concurrently. The selection of tumor protein
 15 markers may be based on routine experiments to determine combinations that results in optimal sensitivity. In addition, or alternatively, assays for tumor proteins provided herein may be combined with assays for other known tumor antigens.

DIAGNOSTIC KITS

The present invention further provides kits for use within any of the above
 20 diagnostic methods. Such kits typically comprise two or more components necessary for performing a diagnostic assay. Components may be compounds, reagents, containers and/or equipment. For example, one container within a kit may contain a monoclonal antibody or fragment thereof that specifically binds to a lung tumor protein. Such antibodies or fragments may be provided attached to a support material, as described
 25 above. One or more additional containers may enclose elements, such as reagents or buffers, to be used in the assay. Such kits may also, or alternatively, contain a detection reagent as described above that contains a reporter group suitable for direct or indirect detection of antibody binding.

Alternatively, a kit may be designed to detect the level of mRNA encoding a lung tumor protein in a biological sample. Such kits generally comprise at least one oligonucleotide probe or primer, as described above, that hybridizes to a polynucleotide encoding a lung tumor protein. Such an oligonucleotide may be used, for example, within

5 a PCR or hybridization assay. Additional components that may be present within such kits include a second oligonucleotide and/or a diagnostic reagent or container to facilitate the detection of a polynucleotide encoding a lung tumor protein.

The following Examples are offered by way of illustration and not by way of limitation.

10

EXAMPLE

EXAMPLE 1

ISOLATION AND CHARACTERIZATION OF cDNA SEQUENCES ENCODING LUNG TUMOR POLYPEPTIDES

5 This example illustrates the isolation of cDNA molecules encoding lung tumor-specific polypeptides from lung tumor cDNA libraries.

A. ISOLATION OF CDNA SEQUENCES FROM A LUNG SQUAMOUS CELL CARCINOMA LIBRARY

10 A human lung squamous cell carcinoma cDNA expression library was constructed from poly A⁺ RNA from a pool of two patient tissues using a Superscript Plasmid System for cDNA Synthesis and Plasmid Cloning kit (BRL Life Technologies, Gaithersburg, MD) following the manufacturer's protocol. Specifically, lung carcinoma tissues were homogenized with polytron (Kinematica, Switzerland) and total RNA was extracted using Trizol reagent (BRL Life Technologies) as directed by the manufacturer.

15 The poly A⁺ RNA was then purified using an oligo dT cellulose column as described in Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratories, Cold Spring Harbor, NY, 1989. First-strand cDNA was synthesized using the NotI/Oligo-dT18 primer. Double-stranded cDNA was synthesized, ligated with BstXI/EcoRI adaptors (Invitrogen, San Diego, CA) and digested with NotI. Following size

20 fractionation with cDNA size fractionation columns (BRL Life Technologies), the cDNA was ligated into the BstXI/NotI site of pcDNA3.1 (Invitrogen) and transformed into ElectroMax *E. coli* DH10B cells (BRL Life Technologies) by electroporation.

 Using the same procedure, a normal human lung cDNA expression library was prepared from a pool of four tissue specimens. The cDNA libraries were characterized

25 by determining the number of independent colonies, the percentage of clones that carried insert, the average insert size and by sequence analysis. The lung squamous cell carcinoma library contained 2.7×10^6 independent colonies, with 100% of clones having an insert and

the average insert size being 2100 base pairs. The normal lung cDNA library contained 1.4×10^6 independent colonies, with 90% of clones having inserts and the average insert size being 1800 base pairs. For both libraries, sequence analysis showed that the majority of clones had a full length cDNA sequence and were synthesized from mRNA

5 cDNA library subtraction was performed using the above lung squamous cell carcinoma and normal lung cDNA libraries, as described by Hara *et al.* (*Blood*, 84:189-199, 1994) with some modifications. Specifically, a lung squamous cell carcinoma-specific subtracted cDNA library was generated as follows. Normal tissue cDNA library (80 μ g) was digested with BamHI and XhoI, followed by a filling-in reaction
10 with DNA polymerase Klenow fragment. After phenol-chloroform extraction and ethanol precipitation, the DNA was dissolved in 133 μ l of H₂O, heat-denatured and mixed with 133 μ l (133 μ g) of Photoprobe biotin (Vector Laboratories, Burlingame, CA). As recommended by the manufacturer, the resulting mixture was irradiated with a 270 W sunlamp on ice for 20 minutes. Additional Photoprobe biotin (67 μ l) was added and the
15 biotinylation reaction was repeated. After extraction with butanol five times, the DNA was ethanol-precipitated and dissolved in 23 μ l H₂O to form the driver DNA.

To form the tracer DNA, 10 μ g lung squamous cell carcinoma cDNA library was digested with NotI and SpeI, phenol chloroform extracted and passed through Chroma spin-400 columns (Clontech, Palo Alto, CA). Typically, 5 μ g of cDNA was recovered
20 after the sizing column. Following ethanol precipitation, the tracer DNA was dissolved in 5 μ l H₂O. Tracer DNA was mixed with 15 μ l driver DNA and 20 μ l of 2 x hybridization buffer (1.5 M NaCl/10 mM EDTA/50 mM HEPES pH 7.5/0.2% sodium dodecyl sulfate), overlaid with mineral oil, and heat-denatured completely. The sample was immediately transferred into a 68 °C water bath and incubated for 20 hours (long hybridization [LH]).
25 The reaction mixture was then subjected to a streptavidin treatment followed by phenol/chloroform extraction. This process was repeated three more times. Subtracted DNA was precipitated, dissolved in 12 μ l H₂O, mixed with 8 μ l driver DNA and 20 μ l of 2 x hybridization buffer, and subjected to a hybridization at 68 °C for 2 hours (short hybridization [SH]). After removal of biotinylated double-stranded DNA, subtracted

cDNA was ligated into NotI/SpeI site of chloramphenicol resistant pBCSK⁺ (Stratagene, La Jolla, CA) and transformed into ElectroMax *E. coli* DH10B cells by electroporation to generate a lung squamous cell carcinoma specific subtracted cDNA library (herein after referred to as “lung subtraction I”).

5 A second lung squamous cell carcinoma specific subtracted cDNA library (referred to as “lung subtraction II”) was generated in a similar way to the lung subtraction library I, except that eight frequently recovered genes from lung subtraction I were included in the driver DNA, and 24,000 independent clones were recovered.

 To analyze the subtracted cDNA libraries, plasmid DNA was prepared from
 10 320 independent clones, randomly picked from the subtracted lung squamous cell carcinoma specific libraries. Representative cDNA clones were further characterized by DNA sequencing with a Perkin Elmer/Applied Biosystems Division Automated Sequencer Model 373A and/or Model 377 (Foster City, CA). The cDNA sequences for sixty isolated clones are provided in SEQ ID NO: 1-60. These sequences were compared to known
 15 sequences in the gene bank using the EMBL and GenBank databases (release 96). No significant homologies were found to the sequences provided in SEQ ID NO: 2, 3, 19, 38 and 46. The sequences of SEQ ID NO: 1, 6-8, 10-13, 15, 17, 18, 20-27, 29, 30, 32, 34-37, 39-45, 47-49, 51, 52, 54, 55 and 57-59 were found to show some homology to previously identified expressed sequence tags (ESTs). The sequences of SEQ ID NO: 9, 28, 31 and 33
 20 were found to show some homology to previously identified non-human gene sequences and the sequences of SEQ ID NO: 4, 5, 14, 50, 53, 56 and 60 were found to show some homology to gene sequences previously identified in humans.

 The subtraction procedure described above was repeated using the above lung squamous cell carcinoma cDNA library as the tracer DNA, and the above normal lung
 25 tissue cDNA library and a cDNA library from normal liver and heart (constructed from a pool of one sample of each tissue as described above), plus twenty other cDNA clones that were frequently recovered in lung subtractions I and II, as the driver DNA (lung subtraction III). The normal liver and heart cDNA library contained 1.76×10^6 independent colonies, with 100% of clones having inserts and the average insert size being 1600 base

pairs. Ten additional clones were isolated (SEQ ID NO: 61-70). Comparison of these cDNA sequences with those in the gene bank as described above, revealed no significant homologies to the sequences provided in SEQ ID NO: 62 and 67. The sequences of SEQ ID NO: 61, 63-66, 68 and 69 were found to show some homology to previously isolated ESTs and the sequence provided in SEQ ID NO: 70 was found to show some homology to a previously identified rat gene.

In further studies, the subtraction procedure described above was repeated using the above lung squamous cell carcinoma cDNA library as the tracer DNA, and a cDNA library from a pool of normal lung, kidney, colon, pancreas, brain, resting PBMC, heart, skin and esophagus as the driver DNA, with esophagus cDNAs making up one third of the driver material. Since esophagus is enriched in normal epithelial cells, including differentiated squamous cells, this procedure is likely to enrich genes that are tumor specific rather than tissues specific. The cDNA sequences of 48 clones determined in this subtraction are provided in SEQ ID NO: 177-224. The sequences of SEQ ID NO: 177, 178, 180, 181, 183, 187, 192, 195-197, 208, 211, 212, 215, 216, 218 and 219 showed some homology to previously identified genes. The sequences of SEQ ID NO: 179, 182, 184-186, 188-191, 193, 194, 198-207, 209 210, 213, 214, 217, 220 and 224 showed some homology to previously determined ESTs. The sequence of SEQ ID NO: 221-223 showed no homology to any previously determined sequence.

B. ISOLATION OF cDNA SEQUENCES FROM A LUNG ADENOCARCINOMA LIBRARY

A human lung adenocarcinoma cDNA expression library was constructed as described above. The library contained 3.2×10^6 independent colonies, with 100% of clones having an insert and the average insert size being 1500 base pairs. Library subtraction was performed as described above using the normal lung and normal liver and heart cDNA expression libraries described above as the driver DNA. Twenty-six hundred independent clones were recovered.

Initial cDNA sequence analysis from 100 independent clones revealed many ribosomal protein genes. The cDNA sequences for fifteen clones isolated in this subtraction are provided in SEQ ID NO: 71-86. Comparison of these sequences with those in the gene bank as described above revealed no significant homologies to the sequence provided in SEQ ID NO: 84. The sequences of SEQ ID NO: 71, 73, 74, 77, 78 and 80-82 were found to show some homology to previously isolated ESTs, and the sequences of SEQ ID NO: 72, 75, 76, 79, 83 and 85 were found to show some homology to previously identified human genes.

In further studies, a cDNA library (referred to as mets3616A) was constructed from a metastatic lung adenocarcinoma. The determined cDNA sequences of 25 clones sequenced at random from this library are provided in SEQ ID NO: 255-279. The mets3616A cDNA library was subtracted against a cDNA library prepared from a pool of normal lung, liver, pancreas, skin, kidney, brain and resting PBMC. To increase the specificity of the subtraction, the driver was spiked with genes that were determined to be most abundant in the mets3616A cDNA library, such as EF1-alpha, integrin-beta and anticoagulant protein PP4, as well as with cDNAs that were previously found to be differentially expressed in subtracted lung adenocarcinoma cDNA libraries. The determined cDNA sequences of 51 clones isolated from the subtracted library (referred to as mets3616A-S1) are provided in SEQ ID NO: 280-330.

Comparison of the sequences of SEQ ID NO: 255-330 with those in the public databases revealed no significant homologies to the sequences of SEQ ID NO: 255-258, 260, 262-264, 270, 272, 275, 276, 279, 281, 287, 291, 296, 300 and 310. The sequences of SEQ ID NO: 259, 261, 265-269, 271, 273, 274, 277, 278, 282-285, 288-290, 292, 294, 297-299, 301, 303-309, 313, 314, 316, 320-324 and 326-330 showed some homology to previously identified gene sequences, while the sequences of SEQ ID NO: 280, 286, 293, 302, 310, 312, 315, 317-319 and 325 showed some homology to previously isolated expressed sequence tags (ESTs).

EXAMPLE 2

DETERMINATION OF TISSUE SPECIFICITY OF LUNG TUMOR POLYPEPTIDES

Using gene specific primers, mRNA expression levels for seven representative lung tumor polypeptides described in Example 1 were examined in a variety of normal and tumor tissues using RT-PCR.

Briefly, total RNA was extracted from a variety of normal and tumor tissues using Trizol reagent as described above. First strand synthesis was carried out using 2 µg of total RNA with SuperScript II reverse transcriptase (BRL Life Technologies) at 42 °C for one hour. The cDNA was then amplified by PCR with gene-specific primers. To ensure the semi-quantitative nature of the RT-PCR, β-actin was used as an internal control for each of the tissues examined. 1 µl of 1:30 dilution of cDNA was employed to enable the linear range amplification of the β-actin template and was sensitive enough to reflect the differences in the initial copy numbers. Using these conditions, the β-actin levels were determined for each reverse transcription reaction from each tissue. DNA contamination was minimized by DNase treatment and by assuring a negative PCR result when using first strand cDNA that was prepared without adding reverse transcriptase.

mRNA Expression levels were examined in five different types of tumor tissue (lung squamous cell carcinoma from 3 patients, lung adenocarcinoma, colon tumor from 2 patients, breast tumor and prostate tumor), and thirteen different normal tissues (lung from 4 donors, prostate, brain, kidney, liver, ovary, skeletal muscle, skin, small intestine, stomach, myocardium, retina and testes). Using a 10-fold amount of cDNA, the antigen LST-S1-90 (SEQ ID NO: 3) was found to be expressed at high levels in lung squamous cell carcinoma and in breast tumor, and at low to undetectable levels in the other tissues examined.

The antigen LST-S2-68 (SEQ ID NO: 15) appears to be specific to lung and breast tumor, however, expression was also detected in normal kidney. Antigens LST-S1-169 (SEQ ID NO: 6) and LST-S1-133 (SEQ ID NO: 5) appear to be very abundant in lung tissues (both normal and tumor), with the expression of these two genes being decreased in most of the normal tissues tested. Both LST-S1-169 and LST-S1-133 were also expressed

in breast and colon tumors. Antigens LST-S1-6 (SEQ ID NO: 7) and LST-S2-I2-5F (SEQ ID NO: 47) did not show tumor or tissue specific expression, with the expression of LST-S1-28 being rare and only detectable in a few tissues. The antigen LST-S3-7 (SEQ ID NO: 63) showed lung and breast tumor specific expression, with its message only being detected in normal testes when the PCR was performed for 30 cycles. Lower level expression was detected in some normal tissues when the cycle number was increased to 35. Antigen LST-S3-13 (SEQ ID NO: 66) was found to be expressed in 3 out of 4 lung tumors, one breast tumor and both colon tumor samples. Its expression in normal tissues was lower compared to tumors, and was only detected in 1 out of 4 normal lung tissues and in normal tissues from kidney, ovary and retina. Expression of antigens LST-S3-4 (SEQ ID NO: 62) and LST-S3-14 (SEQ ID NO: 67) was rare and did not show any tissue or tumor specificity. Consistent with Northern blot analyses, the RT-PCT results on antigen LAT-S1-A-10A (SEQ ID NO: 78) suggested that its expression is high in lung, colon, stomach and small intestine tissues, including lung and colon tumors, whereas its expression was low or undetectable in other tissues.

A total of 2002 cDNA fragments isolated in lung subtractions I, II and III, described above, were colony PCR amplified and their mRNA expression levels in lung tumor, normal lung, and various other normal and tumor tissues were determined using microarray technology (Synteni, Palo Alto, CA). Briefly, the PCR amplification products were dotted onto slides in an array format, with each product occupying a unique location in the array. mRNA was extracted from the tissue sample to be tested, reverse transcribed, and fluorescent-labeled cDNA probes were generated. The microarrays were probed with the labeled cDNA probes, the slides scanned and fluorescence intensity was measured. This intensity correlates with the hybridization intensity. Seventeen non-redundant cDNA clones showed over-expression in lung squamous tumors, with expression in normal tissues tested (lung, skin, lymph node, colon, liver, pancreas, breast, heart, bone marrow, large intestine, kidney, stomach, brain, small intestine, bladder and salivary gland) being either undetectable, or 10-fold less compared to lung squamous tumors. The determined partial cDNA sequences for the clone L513S are provided in SEQ ID NO: 87 and 88; those for

L514S are provided in SEQ ID NO: 89 and 90; those for L516S in SEQ ID NO: 91 and 92; that for L517S in SEQ ID NO: 93; that for L519S in SEQ ID NO: 94; those for L520S in SEQ ID NO: 95 and 96; those for L521S in SEQ ID NO: 97 and 98; that for L522S in SEQ ID NO: 99; that for L523S in SEQ ID NO: 100; that for L524S in SEQ ID NO: 101; that
 5 for L525S in SEQ ID NO: 102; that for L526S in SEQ ID NO: 103; that for L527S in SEQ ID NO: 104; that for L528S in SEQ ID NO: 105; that for L529S in SEQ ID NO: 106; and those for L530S in SEQ ID NO: 107 and 108. Additionally, the full-length cDNA sequence for L530S is provided in SEQ ID NO: 151, with the corresponding predicted amino acid sequence being provided in SEQ ID NO: 152. L530S shows homology to a splice variant
 10 of a p53 tumor suppressor homologue, p63. The cDNA sequences of 7 known isoforms of p63 are provided in SEQ ID NO: 331-337, with the corresponding predicted amino acid sequences being provided in SEQ ID NO: 338-344, respectively.

Due to polymorphisms, the clone L531S appears to have two forms. A first determined full-length cDNA sequence for L531S is provided in SEQ ID NO: 109, with
 15 the corresponding predicted amino acid sequence being provided in SEQ ID NO: 110. A second determined full-length cDNA sequence for L531S is provided in SEQ ID NO: 111, with the corresponding predicted amino acid sequence being provided in SEQ ID NO: 112. The sequence of SEQ ID NO: 111 is identical to that of SEQ ID NO: 109, except that it contains a 27 bp insertion. Similarly, L514S also has two alternatively spliced forms; the
 20 first variant cDNA is listed as SEQ ID NO: 153, with the corresponding amino acid sequence being provided in SEQ ID NO: 155. The second variant form of L514S full-length cDNA is provided in SEQ ID NO: 154, with its corresponding amino acid sequence being provided in SEQ ID NO: 156.

Full length cloning for L524S (SEQ ID NO: 101) yielded two variants
 25 (SEQ ID NO: 163 and 164) with the corresponding predicted amino acid sequences of SEQ ID NO: 165 and 166, respectively. Both variants have been shown to encode parathyroid hormone-related peptide.

Attempts to isolate the full-length cDNA for L519S, resulted in the isolation of the extended cDNA sequence provided in SEQ ID NO: 173, which contains a potential

open reading frame. The predicted amino acid sequence encoded by the sequence of SEQ ID NO: 173 is provided in SEQ ID NO: 174. Additionally, the full-length cDNA sequence for the clone of SEQ ID NO: 100 (known as L523S), a known gene, is provided in SEQ ID NO: 175, with the corresponding predicted amino acid sequence provided in SEQ ID NO: 176. In further studies, a full-length cDNA sequence for L523S was isolated from a L523S-positive tumor cDNA library by PCR amplification using gene specific primers designed from the sequence of SEQ ID NO: 175. The determined cDNA sequence is provided in SEQ ID NO: 347. The amino acid sequence encoded by this sequence is provided in SEQ ID NO: 348. This protein sequence differs from the previously published protein sequence at two amino acid positions, namely at positions 158 and 410.

Comparison of the sequences of L514S and L531S (SEQ ID NO: 87 and 88, 89 and 90, and 109, respectively) with those in the gene bank, as described above, revealed no significant homologies to known sequences. The sequences of L513S, L516S, L517S, L519S, L520S and L530S (SEQ ID NO: 87 and 88, 91 and 92, 93, 94, 95 and 96, 107 and 108, respectively) were found to show some homology to previously identified ESTs. The sequences of L521S, L522S, L523S, L524S, L525S, L526S, L527S, L528S and L529S (SEQ ID NO: 97 and 98, 99, 99, 101, 102, 103, 104, 105, and 106, respectively) were found to represent known genes. The determined full-length cDNA sequences for L520S is provided in SEQ ID NO: 113, with the corresponding predicted amino acid sequence being provided in SEQ ID NO: 114. Subsequent microarray analysis has shown L520S to be overexpressed in breast tumors in addition to lung squamous tumors.

Further analysis has demonstrated that L529S (SEQ ID NO: 106 and 115), L525S (SEQ ID NO: 102 and 120) and L527S (SEQ ID NO: 104) are cytoskeletal components and potentially squamous cell specific proteins. L529S is connexin 26, a gap junction protein. It is highly expressed in lung squamous tumor 9688T, and moderately over-expressed in two others. However, lower level expression of connexin 26 is also detectable in normal skin, colon, liver and stomach. The over-expression of connexin 26 in some breast tumors has been reported and a mutated form of L529S may result in over-expression in lung tumors. L525S is plakophilin 1, a desmosomal protein found in plaque-

bearing adhering junctions of the skin. Expression levels for L525S mRNA is highly elevated in three out of four lung squamous tumors tested, and in normal skin. L527S has been identified as keratin 6 isoform, type II 58 Kd keratin, and cytokeratin 13 and shows over-expression in squamous tumors and low expression in normal skin, breast and colon tissues. Notably, keratin and keratin-related genes have been extensively documented as potential markers for lung cancer including CYFRA2.1 (Pastor, A., et al, *Eur. Respir. J.*, 10:603-609, 1997). L513S (SEQ ID NO: 87 and 88) shows moderate over-expression in several tumor tissues tested, and encodes a protein that was first isolated as a pemphigus vulgaris antigen.

L520S (SEQ ID NO: 95 and 96) and L521S (SEQ ID NO: 97 and 98) are highly expressed in lung squamous tumors, and L520S is up-regulated in normal salivary gland and L521S is over-expressed in normal skin. Both belong to a family of small proline rich proteins and represent markers for fully differentiated squamous cells. L521S has been described as a specific marker for lung squamous tumor (Hu, R., et al, *Lung Cancer*, 20:25-30, 1998). L515S (SEQ ID NO: 162) encodes IGF- β 2 and L516S is an aldose reductase homologue and both are moderately expressed in lung squamous tumors and in normal colon. Notably, L516S (SEQ ID NO: 91 and 92) is up-regulated in metastatic tumors but not primary lung adenocarcinoma, an indication of its potential role in metatasis and a potential prognostic marker. L522S (SEQ ID NO: 99) is moderately over-expressed in lung squamous tumors with minimum expression in normal tissues. L522S has been shown to belong to a class IV alcohol dehydrogenase, ADH7, and its expression profile suggests it is a squamous cell specific antigen. L523S (SEQ ID NO: 100) is moderately over-expressed in lung squamous tumor, human pancreatic cancer cell lines and pancreatic cancer tissues, suggesting this gene may be a shared antigen between pancreatic and lung squamous cell cancer.

L524S (SEQ ID NO: 101) is over-expressed in the majority of squamous tumors tested and is homologous with parathyroid hormone-related peptide (PTHrP), which is best known to cause humoral hypercalcaemia associated with malignant tumors such as leukemia, prostate and breast cancer. It is also believed that PTHrP is most

commonly associated with squamous carcinoma of lung and rarely with lung adenocarcinoma (Davidson, L.A., et al, *J. Pathol.*, 178: 398-401, 1996). L528S (SEQ ID NO: 105) is highly over-expressed in two lung squamous tumors with moderate expression in two other squamous tumors, one lung adenocarcinoma and some normal tissues, including skin, lymph nodes, heart, stomach and lung. It encodes the NMB gene that is similar to the precursor of melanocyte specific gene Pmel17, which is reported to be preferentially expressed in low-metastatic potential melanoma cell lines. This suggests that L528S may be a shared antigen in both melanoma and lung squamous cell carcinoma. L526S (SEQ ID NO: 103) is overexpressed in all lung squamous cell tumor tissues tested and has been shown to share homology with a gene (ATM) in which a mutation causes ataxia telangiectasia, a genetic disorder in humans causing a predisposition to cancer, among other symptoms. ATM encodes a protein that activates p53 mediated cell-cycle checkpoint through direct binding and phosphorylation of the p53 molecule. Approximately 40% of lung cancer is associated with p53 mutations, and it is speculated that over-expression of ATM is a result of compensation for loss of p53 function, but it is unknown whether over-expression is the cause of result of lung squamous cell carcinoma. Additionally, expression of L526S (ATM) is also detected in a metastatic but not lung adenocarcinoma, suggesting a role in metastasis.

Expression of L523S (SEQ ID NO: 175), was also examined by real time RT-PCR as described above. In a first study using a panel of lung squamous tumors, L523S was found to be expressed in 4/7 lung squamous tumors, 2/3 head and neck squamous tumors and 2/2 lung adenocarcinomas, with low level expression being observed in skeletal muscle, soft palate and tonsil. In a second study using a lung adenocarcinoma panel, expression of L523S was observed in 4/9 primary adenocarcinomas, 2/2 lung pleural effusions, 1/1 metastatic lung adenocarcinomas and 2/2 lung squamous tumors, with little expression being observed in normal tissues.

Expression of L523S in lung tumors and various normal tissues was also examined by Northern blot analysis, using standard techniques. In a first study, L523S was found to be expressed in a number of lung adenocarcinomas and squamous cell

carcinomas, as well as normal tonsil. No expression was observed in normal lung. In a second study using a normal tissue blot (HB-12) from Clontech, no expression was observed in brain, skeletal muscle, colon, thymus, spleen, kidney, liver, small intestine, lung or PBMC, although there was strong expression in placenta.

5

EXAMPLE 3

ISOLATION AND CHARACTERIZATION OF LUNG TUMOR POLYPEPTIDES BY
PCR-BASED SUBTRACTION

Eight hundred and fifty seven clones from a cDNA subtraction library, containing cDNA from a pool of two human lung squamous tumors subtracted against
10 eight normal human tissue cDNAs including lung, PBMC, brain, heart, kidney, liver, pancreas, and skin, (Clontech, Palo Alto, CA) were derived and submitted to a first round of PCR amplification. This library was subjected to a second round of PCR amplification, following the manufacturer's protocol. The resulting cDNA fragments were subcloned into the vector P7- Adv vector (Clontech, Palo Alto, CA) and transformed into DH5 α *E. coli*
15 (Gibco, BRL). DNA was isolated from independent clones and sequenced using a Perkin Elmer/Applied Biosystems Division Automated Sequencer Model 373A.

One hundred and sixty two positive clones were sequenced. Comparison of the DNA sequences of these clones with those in the EMBL and GenBank databases, as described above, revealed no significant homologies to 13 of these clones, hereinafter
20 referred to as Contigs 13, 16, 17, 19, 22, 24, 29, 47, 49, 56-59. The determined cDNA sequences for these clones are provided in SEQ ID NO: 125, 127-129, 131-133, 142, 144, 148-150, and 157, respectively. Contigs 1, 3-5, 7-10, 12, 11, 15, 20, 31, 33, 38, 39, 41, 43, 44, 45, 48, 50, 53, 54 (SEQ ID NO: 115-124, 126, 130, 134-141, 143, 145-147, respectively) were found to show some degree of homology to previously identified DNA
25 sequences. Contig 57 (SEQ ID NO: 149) was found to represent the clone L519S (SEQ ID NO: 94) disclosed in US. Patent Application No. 09/123,912, filed July 27, 1998. To the best of the inventors' knowledge, none of these sequences have been previously shown to be differentially over-expressed in lung tumors.

mRNA expression levels for representative clones in lung tumor tissues, normal lung tissues (n=4), resting PBMC, salivary gland, heart, stomach, lymph nodes, skeletal muscle, soft palate, small intestine, large intestine, bronchial, bladder, tonsil, kidney, esophagus, bone marrow, colon, adrenal gland, pancreas, and skin, (all derived from human) were determined by RT-PCR as described above. Expression levels using microarray technology, as described above, were examined in one sample of each tissue type unless otherwise indicated.

Contig 3 (SEQ ID NO: 116) was found to be highly expressed in all head and neck squamous cell tumors tested (17/17), and expressed in the majority (8/12) of lung squamous tumors, (high expression in 7/12, moderate in 2/12, and low in 2/12), while showing negative expression for 2/4 normal lung tissues and low expression in the remaining two samples. Contig 3 showed moderate expression in skin and soft palate, and lowered expression levels in resting PBMC, large intestine, salivary gland, tonsil, pancreas, esophagus, and colon. Contig 11 (SEQ ID NO: 124) was found to be expressed in all head and neck squamous cell tumors tested (17/17): highly expressed in 14/17, and moderately expressed in 3/17. Additionally, expression in lung squamous tumors showed high expression in 3/12 and moderate in 4/12. Contig 11 was negative for 3/4 normal lung samples, with the remaining sample having only low expression. Contig 11 showed low to moderate reactivity to salivary gland, soft palate, bladder, tonsil, skin, esophagus, and large intestine. Contig 13 (SEQ ID NO: 125) was found to be expressed in all head and neck squamous cell tumors tested (17/17): highly expressed in 12/17, and moderately expressed in 5/17. Contig 13 was expressed in 7/12 lung squamous tumors, with high expression in 4/12 and moderate expression in three samples. Analysis of normal lung samples showed negative expression for 2/4 and low to moderate expression in the remaining two samples. Contig 13 did show low to moderate reactivity to resting PBMC, salivary gland, bladder, pancreas, tonsil, skin, esophagus, and large intestine, as well as high expression in soft palate. Subsequent full-length cloning efforts revealed that contig 13 (also known as L761P) maps to the 3' untranslated region of the hSec10p gene. The full-length sequence for this gene is set forth in SEQ ID NO: 368, and encodes the protein set forth in SEQ ID

NO: 369. Contig 16 (SEQ ID NO: 127) was found to be moderately expressed in some head and neck squamous cell tumors (6/17) and one lung squamous tumor; while showing no expression in any normal lung samples tested. Contig 16 did show low reactivity to resting PBMC, large intestine, skin, salivary gland, and soft palate. Contig 17 (SEQ ID NO: 128) was shown to be expressed in all head and neck squamous cell tumors tested (17/17): highly expressed in 5/17, and moderately expressed in 12/17. Expression levels in lung squamous tumors showed one tumor sample with high expression and 3/12 with moderate levels. Contig 17 was negative for 2/4 normal lung samples, with the remaining samples having only low expression. Additionally, low level expression was found in esophagus and soft palate. Contig 19 (SEQ ID NO: 129) was found to be expressed in most head and neck squamous cell tumors tested (11/17); with two samples having high levels, 6/17 showing moderate expression, and low expression being found in 3/17. Testing in lung squamous tumors revealed only moderate expression in 3/12 samples. Expression levels in 2/4 of normal lung samples were negative, the two other samples having only low expression. Contig 19 showed low expression levels in esophagus, resting PBMC, salivary gland, bladder, soft palate and pancreas.

Contig 22 (SEQ ID NO: 131), was shown to be expressed in most head and neck squamous cell tumors tested (13/17) with high expression in four of these samples, moderate expression in 6/17, and low expression in 3/17. Expression levels in lung squamous tumors were found to be moderate to high for 3/12 tissues tested, with negative expression in two normal lung samples and low expression in two other samples (n=4). Contig 22 showed low expression in skin, salivary gland and soft palate. Similarly, Contig 24 (SEQ ID NO: 132) was found to be expressed in most head and neck squamous cell tumors tested (13/17) with high expression in three of these samples, moderate expression in 6/17, and low expression in 4/17. Expression levels in lung squamous tumors were found to be moderate to high for 3/12 tissues tested, with negative expression for three normal lung samples and low expression in one sample (n=4). Contig 24 showed low expression in skin, salivary gland and soft palate. Contig 29 (SEQ ID NO: 133) was expressed in nearly all head and neck squamous cell tumors tested (16/17): highly

expressed in 4/17, moderately expressed in 11/17, with low expression in one sample. Also, it was moderately expressed in 3/12 lung squamous tumors, while being negative for 2/4 normal lung samples. Contig 29 showed low to moderate expression in large intestine, skin, salivary gland, pancreas, tonsil, heart and soft palate. Contig 47 (SEQ ID NO: 142) was expressed in most head and neck squamous cell tumors tested (12/17): moderate expression in 10/17, and low expression in two samples. In lung squamous tumors, it was highly expressed in one sample and moderately expressed in two others (n=13). Contig 47 was negative for 2/4 normal lung samples, with the remaining two samples having moderate expression. Also, Contig 47 showed moderate expression in large intestine, and pancreas, and low expression in skin, salivary gland, soft palate, stomach, bladder, resting PBMC, and tonsil.

Contig 48 (SEQ ID NO: 143) was expressed in all head and neck squamous cell tumors tested (17/17): highly expressed in 8/17 and moderately expressed in 7/17, with low expression in two samples. Expression levels in lung squamous tumors were high to moderate in three samples (n=13). Contig 48 was negative for one out of four normal lung samples, the remaining showing low or moderate expression. Contig 48 showed moderate expression in soft palate, large intestine, pancreas, and bladder, and low expression in esophagus, salivary gland, resting PBMC, and heart. Contig 49 (SEQ ID NO: 144) was expressed at low to moderate levels in 6/17 head and neck squamous cell tumors tested. Expression levels in lung squamous tumors were moderate in three samples (n=13). Contig 49 was negative for 2/4 normal lung samples, the remaining samples showing low expression. Moderate expression levels in skin, salivary gland, large intestine, pancreas, bladder and resting PBMC were shown, as well as low expression in soft palate, lymph nodes, and tonsil. Contig 56 (SEQ ID NO: 148) was expressed in low to moderate levels in 3/17 head and neck squamous cell tumors tested, and in lung squamous tumors, showing low to moderate levels in three out of thirteen samples. Notably, low expression levels were detected in one adenocarcinoma lung tumor sample (n=2). Contig 56 was negative for 3/4 normal lung samples, and showed moderate expression levels in only large intestine, and low expression in salivary gland, soft palate, pancreas, bladder, and resting

PBMC. Contig 58, also known as L769P, (SEQ ID NO: 150) was expressed at moderate levels in 11/17 head and neck squamous cell tumors tested and low expression in one additional sample. Expression in lung squamous tumors showed low to moderate levels in three out of thirteen samples. Contig 58 was negative for 3/4 normal lung samples, with one sample having low expression. Moderate expression levels in skin, large intestine, and resting PBMC were demonstrated, as well as low expression in salivary gland, soft palate, pancreas, and bladder. Contig 59 (SEQ ID NO: 157) was expressed in some head, neck, and lung squamous tumors. Low level expression of Contig 59 was also detected in salivary gland and large intestine.

The full-length cDNA sequence for Contig 22, also referred to as L763P, is provided in SEQ ID NO: 158, with the corresponding predicted amino acid sequence being provided in SEQ ID NO: 159. Real-time RT-PCR analysis of L763P revealed that it is highly expressed in 3/4 lung squamous tumors as well as 4/4 head and neck squamous tumors, with low level expression being observed in normal brain, skin, soft pallet and trachea. Subsequent database searches revealed that the sequence of SEQ ID NO: 158 contains a mutation, resulting in a frameshift in the corresponding protein sequence. A second cDNA sequence for L763P is provided in SEQ ID NO: 345, with the corresponding amino acid sequence being provided in SEQ ID NO: 346. The sequences of SEQ ID NO: 159 and 346 are identical with the exception of the C-terminal 33 amino acids of SEQ ID NO: 159.

The full-length cDNA sequence incorporating Contigs 17, 19, and 24, referred to as L762P, is provided in SEQ ID NO: 160, with the corresponding predicted amino acid sequence being provided in SEQ ID NO: 161. Further analysis of L762P has determined it to be a type I membrane protein and two additional variants have been sequenced. Variant 1 (SEQ ID NO: 167, with the corresponding amino acid sequence in SEQ ID NO: 169) is an alternatively spliced form of SEQ ID NO: 160 resulting in deletion of 503 nucleotides, as well as deletion of a short segment of the expressed protein. Variant 2 (SEQ ID NO: 168, with the corresponding amino acid sequence in SEQ ID NO: 170) has a two nucleotide deletion at the 3' coding region in comparison to SEQ ID NO: 160,

resulting in a secreted form of the expressed protein. Real-time RT-PCR analysis of L762P revealed that is over-expressed in 3/4 lung squamous tumors and 4/4 head & neck tumors, with low level expression being observed in normal skin, soft pallet and trachea.

An epitope of L762 was identified as having the sequence
 5 KPGHWTYTLNNTTHSLQALK, amino acids 571-590 of SEQ ID NO:161.

The full-length cDNA sequence for contig 56 (SEQ ID NO: 148), also referred to as L773P, is provided in SEQ ID NO: 171, with the predicted amino acid sequence in SEQ ID NO: 172. L773P was found to be identical to dihydroxyl dehydrogenase at the 3' portion of the gene, with divergent 5' sequence. As a result, the 69
 10 N-terminal amino acids are unique. The cDNA sequence encoding the 69 N-terminal amino acids is provided in SEQ ID NO: 349, with the N-terminal amino acid sequence being provided in SEQ ID NO: 350. Real-time PCR revealed that L773P is highly expressed in lung squamous tumor and lung adenocarcinoma, with no detectable expression in normal tissues. Subsequent Northern blot analysis of L773P demonstrated
 15 that this transcript is differentially over-expressed in squamous tumors and detected at approximately 1.6 Kb in primary lung tumor tissue and approximately 1.3 Kb in primary head and neck tumor tissue.

Subsequent microarray analysis has shown Contig 58, also referred to as L769S (SEQ ID NO: 150), to be overexpressed in breast tumors in addition to lung
 20 squamous tumors.

EXAMPLE 4

SYNTHESIS OF POLYPEPTIDES

Polypeptides may be synthesized on a Perkin Elmer/Applied Biosystems Division 430A peptide synthesizer using FMOC chemistry with HPTU (O-Benzotriazole-
 25 N,N,N',N'-tetramethyluronium hexafluorophosphate) activation. A Gly-Cys-Gly sequence may be attached to the amino terminus of the peptide to provide a method of conjugation, binding to an immobilized surface, or labeling of the peptide. Cleavage of the peptides from the solid support may be carried out using the following cleavage mixture:

trifluoroacetic acid:ethanedithiol:thioanisole:water:phenol (40:1:2:2:3). After cleaving for 2 hours, the peptides may be precipitated in cold methyl-t-butyl-ether. The peptide pellets may then be dissolved in water containing 0.1% trifluoroacetic acid (TFA) and lyophilized prior to purification by C18 reverse phase HPLC. A gradient of 0%-60% acetonitrile (containing 0.1% TFA) in water (containing 0.1% TFA) may be used to elute the peptides. Following lyophilization of the pure fractions, the peptides may be characterized using electrospray or other types of mass spectrometry and by amino acid analysis.

EXAMPLE 5

PREPARATION OF ANTIBODIES AGAINST LUNG CANCER ANTIGENS

Polyclonal antibodies against the lung cancer antigens L514S, L528S, L531S and L523 (SEQ ID NO: 155, 225, 112 and 176 respectively) were prepared as follows.

Rabbits were immunized with recombinant protein expressed in and purified from *E. coli* as described above. For the initial immunization, 400 µg of antigen combined with muramyl dipeptide (MDP) was injected subcutaneously (S.C.). Animals were boosted S.C. 4 weeks later with 200 µg of antigen mixed with incomplete Freund's Adjuvant (IFA). Subsequent boosts of 100 µg of antigen mixed with IFA were injected S.C. as necessary to induce high antibody titer responses. Serum bleeds from immunized rabbits were tested for antigen-specific reactivity using ELISA assays with purified protein. Polyclonal antibodies against L514S, L528S, L531S and L523S were affinity purified from high titer polyclonal sera using purified protein attached to a solid support.

Immunohistochemical analysis using polyclonal antibodies against L514S was performed on a panel of 5 lung tumor samples, 5 normal lung tissue samples and normal colon, kidney, liver, brain and bone marrow. Specifically, tissue samples were fixed in formalin solution for 24 hours and embedded in paraffin before being sliced into 10 micron sections. Tissue sections were permeabilized and incubated with antibody for 1 hr. HRP-labeled anti-mouse followed by incubation with DAB chromogen was used to visualize L514S immunoreactivity. L514S was found to be highly expressed in lung tumor

tissue with little or no expression being observed in normal lung, brain or bone marrow. Light staining was observed in colon (epithelial crypt cells positive) and kidney (tubules positive). Staining was seen in normal liver but no mRNA has been detected in this tissue making this result suspect.

- 5 Using the same procedure, immunohistochemical analysis using polyclonal antibodies against L528S demonstrated staining in lung tumor and normal lung samples, light staining in colon and kidney and no staining in liver and heart.

- Immunohistochemical analysis using polyclonal antibodies against L531S demonstrated staining in lung tumor samples, light membrane staining in most normal lung
10 samples, epithelial staining in colon, tubule staining in kidney, ductal epithelial staining in liver and no staining in heart.

 Immunohistochemical analysis using polyclonal antibodies against L523S demonstrated staining in all lung cancer samples tested but no staining in normal lung, kidney, liver, colon, bone marrow or cerebellum.

- 15 Generation of polyclonal anti-sera against L762P (SEQ ID NO: 169 and 170) was performed as follows. 400 micrograms of lung antigen was combined with 100 micrograms of muramyl dipeptide (MDP). Equal volume of Incomplete Freund's Adjuvant (IFA) was added and then mixed until an emulsion was formed. Rabbits were injected subcutaneously (S.C.). After four weeks the animals were injected S.C. with 200
20 micrograms of antigen that was mixed with an equal volume of IFA. Every four weeks animals were boosted with 100 micrograms of antigen. Seven days following each boost the animal was bled. Sera was generated by incubating the blood at 4°C for 12-24 hours followed by centrifugation.

- Characterization of polyclonal antisera was carried out as follows. 96 well
25 plates were coated with antigen by incubating with 50 microliters (typically 1 microgram) at 4°C for 20 hrs. 250 microliters of BSA blocking buffer was added to the wells and incubated at RT for 2 hrs. Plates were washed 6 times with PBS/0.01% tween. Rabbit sera was diluted in PBS. Fifty microliters of diluted sera was added to each well and incubated at RT for 30 min. Plates were washed as described above before 50 microliters of goat

anti-rabbit horse radish peroxidase (HRP) at a 1:10000 dilution was added and incubated at RT for 30 min. Plates were washed as described above and 100µl of TMB Microwell Peroxidase Substrate was added to each well. Following a 15 minute incubation in the dark at room temperature the colorimetric reaction was stopped with 100µl 1N H₂SO₄ and read immediately at 450 nm. Antisera showed strong reactivity to antigen L762P.

Immunohistochemical analysis using polyclonal antibodies against L762S demonstrated staining in all lung cancer samples tested, some light staining in the bronchiole epithelium of normal lung, tubule staining in kidney, light epithelial staining in colon and no staining in heart or liver.

10

EXAMPLE 6

PEPTIDE PRIMING OF MICE AND PROPAGATION OF CTL LINES

Immunogenic peptides from the lung cancer antigen L762P (SEQ ID NO: 161) for HLA-A2/K^b-restricted CD8⁺ T cells were identified as follows.

15

The location of HLA-A2 binding peptides within the lung cancer antigen L762P (SEQ ID NO: 161) was predicted using a computer program which predicts peptides sequences likely to be to HLA-A*0201 by fitting to the known peptide binding motif for HLA-A*0201 (Rupert *et al.* (1993) *Cell* 74:929; Rammensee *et al.* (1995) *Immunogenetics* 41:178-228). A series of 19 synthetic peptides corresponding to a selected subset of the predicted HLA-A*0201 binding peptides was prepared as described above.

20

Mice expressing the transgene for human HLA A2/K^b (provided by Dr L. Sherman, The Scripps Research Institute, La Jolla, CA) were immunized with the synthetic peptides, as described by Theobald *et al.*, *Proc. Natl. Acad. Sci. USA* 92:11993-11997, 1995 with the following modifications. Mice were immunized with 50µg of L762P peptide and 120µg of an I-A^b binding peptide derived from hepatitis B Virus protein emulsified in incomplete Freund's adjuvant. Three weeks later these mice were sacrificed and single cell suspensions prepared. Cells were then resuspended at 7×10^6 cells/ml in complete media

25

(RPMI-1640; Gibco BRL, Gaithersburg, MD) containing 10% FCS, 2mM Glutamine (Gibco BRL), sodium pyruvate (Gibco BRL), non-essential amino acids (Gibco BRL), 2×10^{-5} M 2-mercaptoethanol, 50U/ml penicillin and streptomycin, and cultured in the presence of irradiated (3000 rads) L762P peptide- (5 μ g/ml) and 10mg/ml B₂-microglobulin- (3 μ g/ml) LPS blasts (A2 transgenic spleens cells cultured in the presence of 7 μ g/ml dextran sulfate and 25 μ g/ml LPS for 3 days). After six days, cells (5×10^5 /ml) were restimulated with 2.5×10^6 /ml peptide pulsed irradiated (20,000 rads) EL4A2Kb cells (Sherman et al, *Science* 258:815-818, 1992) and 5×10^6 /ml irradiated (3000 rads) A2/K^b-transgenic spleen feeder cells. Cells were cultured in the presence of 10U/ml IL-2. Cells were restimulated on a weekly basis as described, in preparation for cloning the line.

Peptide-specific cell lines were cloned by limiting dilution analysis with irradiated (20,000 rads) L762P peptide-pulsed EL4 A2Kb tumor cells (1×10^4 cells/well) as stimulators and irradiated (3000 rads) A2/K^b-transgenic spleen cells as feeders (5×10^5 cells/ well) grown in the presence of 10U/ml IL-2. On day 7, cells were restimulated as before. On day 14, clones that were growing were isolated and maintained in culture.

Cell lines specific for L762P-87 (SEQ ID NO: 226; corresponding to amino acids 87-95 of SEQ ID NO: 161), L762P-145 (SEQ ID NO: 227; corresponding to amino acids 145-153 of SEQ ID NO: 161), L762P-585 (SEQ ID NO: 228; corresponding to amino acids 585-593 of SEQ ID NO: 161), L762P-425 (SEQ ID NO: 229; corresponding to amino acids 425-433 of SEQ ID NO: 161), L762P(10)-424 (SEQ ID NO: 230; corresponding to amino acids 424-433 of SEQ ID NO: 161) and L762P(10)-458 (SEQ ID NO: 231; corresponding to amino acids 458-467 of SEQ ID NO: 161) demonstrated significantly higher reactivity (as measured by percent specific lysis) against L762P peptide-pulsed EL4-A2/K^b tumor target cells than control peptide-pulsed EL4-A2/K^b tumor target cells.

EXAMPLE 7

IDENTIFICATION OF CD4 IMMUNOGENIC T CELL EPITOPES DERIVED FROM
THE LUNG CANCER ANTIGEN L762P

CD4 T cell lines specific for the antigen L762P (SEQ ID NO: 161) were
5 generated as follows.

A series of 28 overlapping peptides were synthesized that spanned
approximately 50% of the L762P sequence. For priming, peptides were combined into
pools of 4-5 peptides, pulsed at 20 micrograms/ml into dendritic cells for 24 hours. The
dendritic cells were then washed and mixed with positively selected CD4+ T cells in 96
10 well U-bottomed plates. Forty cultures were generated for each peptide pool. Cultures
were restimulated weekly with fresh dendritic cells loaded with peptide pools. Following a
total of 3 stimulation cycles, cells were rested for an additional week and tested for
specificity to antigen presenting cells (APC) pulsed with peptide pools using interferon-
gamma ELISA and proliferation assays. For these assays, adherent monocytes loaded with
15 either the relevant peptide pool or an irrelevant peptide were used as APC. T cell lines that
appeared to specifically recognize L762P peptide pools both by cytokine release and
proliferation were identified for each pool. Emphasis was placed on identifying T cells
with proliferative responses. T cell lines that demonstrated either both L762P-specific
cytokine secretion and proliferation, or strong proliferation alone were further expanded to
20 be tested for recognition of individual peptides from the pools, as well as for recognition of
recombinant L762P. The source of recombinant L762P was *E. coli*, and the material was
partially purified and endotoxin positive. These studies employed 10 micrograms of
individual peptides, 10 or 2 micrograms of an irrelevant peptide, and 2 or 0.5 micrograms
of either L762P protein or an irrelevant, equally impure, *E. coli* generated recombinant
25 protein. Significant interferon-gamma production and CD4 T cell proliferation was
induced by a number of L762P-derived peptides in each pool. The amino acid sequences
for these peptides are provided in SEQ ID NO: 232-251. These peptides correspond to
amino acids 661-680, 676-696, 526-545, 874-893, 811-830, 871-891, 856-875, 826-845,

795-815, 736-755, 706-725, 706-725, 691-710, 601-620, 571-590, 556-575, 616-635, 646-665, 631-650, 541-560 and 586-605, respectively, of SEQ ID NO: 161.

CD4 T cell lines that demonstrated specificity for individual L762P-derived peptides were further expanded by stimulation with the relevant peptide at 10 micrograms/ml. Two weeks post-stimulation, T cell lines were tested using both proliferation and IFN-gamma ELISA assays for recognition of the specific peptide. A number of previously identified T cells continued to demonstrate L762P-peptide specific activity. Each of these lines was further expanded on the relevant peptide and, following two weeks of expansion, tested for specific recognition of the L762P-peptide in titration experiments, as well as for recognition of recombinant *E. coli*-derived L762P protein. For these experiments, autologous adherent monocytes were pulsed with either the relevant L762P-derived peptide, an irrelevant mammaglobin-derived peptide, recombinant *E. coli*-derived L762P (approx. 50% pure), or an irrelevant *E. coli*-derived protein. The majority of T cell lines were found to show low affinity for the relevant peptide, since specific proliferation and IFN-gamma ratios dramatically decreased as L762P peptide was diluted. However, four lines were identified that demonstrated significant activity even at 0.1 micrograms/ml peptide. Each of these lines (referred to as A/D5, D/F5, E/A7 and E/B6) also appeared to specifically proliferate in response to the *E. coli*-derived L762P protein preparation, but not in response to the irrelevant protein preparation. The amino acid sequences of the L762P-derived peptides recognized by these lines are provided in SEQ ID NO: 234, 249, 236 and 245, respectively. No protein specific IFN-gamma was detected for any of the lines. Lines A/D5, E/A7 and E/B6 were cloned on autologous adherent monocytes pulsed with the relevant peptide at 0.1 (A/D5 and E/A7) or 1 (D/F5) microgram/ml. Following growth, clones were tested for specificity for the relevant peptide. Numerous clones specific for the relevant peptide were identified for lines A/D5 and E/A7.

EXAMPLE 8

PROTEIN EXPRESSION OF LUNG TUMOR-SPECIFIC ANTIGENS

a) Expression of L514S in *E. coli*

The lung tumor antigen L514S (SEQ ID NO: 89) was subcloned into the expression vector pE32b at NcoI and NotI sites, and transformed into *E. coli* using standard techniques. The protein was expressed from residues 3-153 of SEQ ID NO: 89. The expressed amino acid sequence and the corresponding DNA sequence are provided in SEQ ID NO: 252 and 253, respectively.

b) Expression of L762P

Amino acids 32-944 of the lung tumor antigen L762P (SEQ ID NO: 161), with a 6X His Tag, were subcloned into a modified pET28 expression vector, using kanamycin resistance, and transformed into BL21 CodonPlus using standard techniques. Low to moderate levels of expression were observed. The determined DNA sequence of the L762P expression construct is provided in SEQ ID NO: 254.

EXAMPLE 9

IDENTIFICATION OF MHC CLASS II RESTRICTING ALLELE FOR L-762 PEPTIDE-SPECIFIC RESPONSES

A panel of HLA mismatched antigen presenting cells (APC) were used to identify the MHC class II restricting allele for the L762-peptide specific responses of CD4 T cell clones derived from lines that recognized L762 peptide and recombinant protein. Clones from two lines, AD-5 and EA-7, were tested. The AD-5 derived clones were found to be restricted by the HLA-DRB-1101 allele, and an EA-7 derived clone was found to be restricted by the HLA DRB-0701 or DQB1-0202 allele. Identification of the restriction allele allows targeting of vaccine therapies using the defined peptide to individuals that express the relevant class II allele. Knowing the relevant restricting allele will also enable

clinical monitoring for responses to the defined peptide since only individuals that express the relevant allele will be monitored.

CD4 T cell clones derived from line AD-5 and EA-7 were stimulated on autologous APC pulsed with the specific peptide at 10 µg/ml, and tested for recognition of autologous APC (D72) as well as against a panel of APC partially matched with D72 at class II alleles. Table 1 shows the HLA class typing of the APC tested. Adherent monocytes (generated by 2 hour adherence) from D45, D187, D208, and D326 were used as APC in these experiments. Autologous APC (D72) were not included in the experiment. Each of the APC were pulsed with the relevant peptide (5a for AD-5 and 3e for 3A-7) or the irrelevant mammoglobin peptide at 10 µg/ml, and cultures were established for 10,000 T cells and about 20,000 APC/well. As shown in Table 2, specific proliferation and cytokine production could be detected only when partially matched donor cells were used as APC. Based on the MHC typing analysis, these results strongly suggest that the restricting allele for the L762-specific response of the AD-5 derived clones is HLA-DRB-1101 and for the EA-7 derived clone the restricting allele is HLA DRB-0701 or DQB1-0202.

TABLE 1 - HLA TYPING OF APC

DONOR	DR	DR	DQ	DQ
D72	B1-1101	B1-0701	B1-0202	B1-0301
D45	-3	-15	B1-0201	B1-0602
D187	-4	-15	-1	-7
D208	B1-1101	B1-0407	-3	-3
D326	B1-0301	B1-0701	B1-0202	B1-0201

HLA DR Allele Frequency Data for Donor D72

TABLE 2 - L762 PEPTIDE RESPONSES MAP TO HLA DR ALLELES

	AD-5																								EA-7							
	A11			B10			C10			C11			E6			F1			F9			G8			G9			G10			G12	
	ProI	γ-IFN		ProI	γ-IFN		ProI	γ-IFN		ProI	γ-IFN		ProI	γ-IFN		ProI	γ-IFN		ProI	γ-IFN		ProI	γ-IFN		ProI	γ-IFN		ProI	γ-IFN			
Donor																																
D72 DR-0701, -1101, DQ-0202, -7	46			31			34			24			31			40			55			45			43			91		10		
D45 DR-3,-15, DQ-1, -0201	32	1.7		5.5	1.2		3.3	1		1.0	1.5		1.1	1.1		1.6	1.1		1.4	1.3		0.2	1.1		1.1	1.1		1.2	1.5	0.8	1.1	
D187 DR-4, -15, DQ-1,-7	14	1.2		1.3	1		1.4	1.1		1.4	1.7		1.0	1.1		1.4	1.2		1.2	1.1		0.9	1		1.0	1		1.0	1.6	0.5	1	
D208 DR-4, -1101, DQ-3	138	13		38	54		18.8	10		14.6	4.6		15.3	6.1		45.9	8.6		73.3	14.1		38.0	7.7		174.3	16.1		113.6	19.6	0.8	1	
D326 DR-3, -0701, DQ-0202	0.7	4		0.3	1		0.3	1.4		1.0	2		0.8	1.1		0.3	1.1		0.7	1.1		0.6	1.2		0.4	1		1.2	5	14.1	6.8	

EXAMPLE 10

FUSION PROTEINS OF N-TERMINAL AND C-TERMINAL PORTIONS OF L763P

In another embodiment, a *Mycobacterium tuberculosis*-derived Ra12 polynucleotide is linked to at least an immunogenic portion of a polynucleotide of this invention. Ra12 compositions and methods for their use in enhancing expression of heterologous polynucleotide sequences are described in U.S. Patent Application 60/158,585, the disclosure of which is incorporated herein by reference in its entirety. Briefly, Ra12 refers to a polynucleotide region that is a subsequence of a *Mycobacterium tuberculosis* MTB32A nucleic acid. MTB32A is a serine protease of 32 KD molecular weight encoded by a gene in virulent and avirulent strains of *M. tuberculosis*. The nucleotide sequence and amino acid sequence of MTB32A have been described (for example, U.S. Patent Application 60/158,585; see also, Skeiky *et al.*, *Infection and Immun.* (1999) 67:3998-4007, incorporated herein by reference). Surprisingly, it was discovered that a 14 KD C-terminal fragment of the MTB32A coding sequence expresses at high levels on its own and remains as a soluble protein throughout the purification process. Moreover, Ra12 may enhance the immunogenicity of heterologous antigenic polypeptides with which it is fused. This 14 KD C-terminal fragment of the MTB32A is referred herein as Ra12 and represents a fragment comprising some or all of amino acid residues 192 to 323 of MTB32A.

Recombinant nucleic acids, which encode a fusion polypeptide comprising a Ra12 polypeptide and a heterologous lung tumor polypeptide of interest, can be readily constructed by conventional genetic engineering techniques. Recombinant nucleic acids are constructed so that, preferably, a Ra12 polynucleotide sequence is located 5' to a selected heterologous lung tumor polynucleotide sequence. It may also be appropriate to place a Ra12 polynucleotide sequence 3' to a selected heterologous polynucleotide sequence or to insert a heterologous polynucleotide sequence into a site within a Ra12 polynucleotide sequence.

In addition, any suitable polynucleotide that encodes a Ra12 or a portion or other variant thereof can be used in constructing recombinant fusion polynucleotides comprising Ra12 and one or more lung tumor polynucleotides disclosed herein. Preferred Ra12 polynucleotides generally comprise at least about 15 consecutive nucleotides, at least
 5 about 30 nucleotides, at least about 60 nucleotides, at least about 100 nucleotides, at least about 200 nucleotides, or at least about 300 nucleotides that encode a portion of a Ra12 polypeptide.

Ra12 polynucleotides may comprise a native sequence (*i.e.*, an endogenous sequence that encodes a Ra12 polypeptide or a portion thereof) or may comprise a variant
 10 of such a sequence. Ra12 polynucleotide variants may contain one or more substitutions, additions, deletions and/or insertions such that the biological activity of the encoded fusion polypeptide is not substantially diminished, relative to a fusion polypeptide comprising a native Ra12 polypeptide. Variants preferably exhibit at least about 70% identity, more preferably at least about 80% identity and most preferably at least about 90% identity to a
 15 polynucleotide sequence that encodes a native Ra12 polypeptide or a portion thereof.

Two specific embodiments of fusions between Ra12 and antigens of the present invention are described in this example.

A. N-Terminal Portion of L763P

A fusion protein of full-length Ra12 and the N-terminal portion of L763P
 20 (amino acid residues 1-130) was expressed as a single recombinant protein in *E. coli*. The cDNA for the N-terminal portion was obtained by PCR with a cDNA for the full length L763P and primers L763F3 5' CGGCGAATTCAT-GGATTGGGGGACGCTGC and 1763RV3 5' CGGCCTCGAGTCACCCCTCTA-TCCGAACCTTCTGC. The PCR product with expected size was recovered from agarose gel, digested with restriction
 25 enzymes EcoRI and XhoI, and cloned into the corresponding sites in the expression vector pCRX1. The sequence for the fusion of full-length of Ra12 and L763P-N was confirmed by DNA sequencing (SEQ ID NO:351 and 352).

B. C-Terminal Portion of L763P

A fusion protein of full-length Ra12 and the C-terminal portion of L763P (amino acid residues 100-262) was expressed as a single recombinant protein in *E. coli*. The cDNA of the C-terminal portion of L763P was obtained by PCR with a cDNA for the full length of L763P and primers L763F4 5' CGGCGAATTCCACGAACCACTCGCAAGTTCAG and L763RV4 5' CGGCTCGAG-TTAGCTTGGGCCTGTGATTGC. The PCR product with expected size was recovered from agarose gel, digested with restriction enzymes EcoRI and XhoI, and cloned into the corresponding sites in the expression vector pCRX1. The sequence for the fusion of full-length Ra12 and L763P-C was confirmed by DNA sequencing (SEQ ID NO:353 and 354).

The recombinant proteins described in this example are useful for the preparation of vaccines, for antibody therapeutics, and for diagnosis of lung tumors.

EXAMPLE 11

15 EXPRESSION IN *E. COLI* OF L762P HIS TAG FUSION PROTEIN

PCR was performed on L762P coding region with the following primers:

Forward Primer starting at amino acid 32.

PDM-278 5'ggagtacagcttcaagacaatggg 3' (SEQ ID NO:355) Tm 57°C.

Reverse Primer including natural stop codon after amino acid 920, creating
20 EcoRI site

PDM-280 5'ccatgggaattcattataataattttgtcc 3' (SEQ ID NO:356) TM55°C.

The PCR product was then digested with EcoRI restriction enzyme, gel purified and then cloned into pPDM His, a modified pET28 vector with a His tag in frame, which had been digested with Eco72I and EcoRI restriction enzymes. The correct
25 construct was confirmed by DNA sequence analysis and then transformed into BL21 (DE3) pLys S and BL21 (DE3) CodonPlus RIL expression hosts.

The protein sequence of expressed recombinant L762P is shown in SEQ ID NO:357, and the DNA sequence is shown in SEQ ID NO:358.

EXAMPLE 12

EXPRESSION IN E. COLI OF L773P A, HIS TAG FUSION PROTEIN

The L773P A coding region was PCR amplified using the following primers:

- 5 Forward primer for L773P A starting at amino acid 2.
 PDM-299 5'tggcagccccctcttcttcaagtggc 3' (SEQ ID NO:359) Tm63°C.
 Reverse primer for L773P A creating artificial stop codon after amino acid
 70.
 PDM-355 5'cgccagaattcatcaacaaatctgtagcacc 3' (SEQ ID NO:360)
 10 Tm62°C.

The PCR product was then digested with EcoRI restriction enzyme, gel purified and then cloned into pPDM His, a modified pET28 vector with a His tag in frame, which had been digested with Eco72I and EcoRI restriction enzymes. The correct construct was confirmed by DNA sequence analysis and then transformed into BL21 (DE3) pLys S and BL21 (DE3) CodonPlus RIL expression hosts.

The protein sequence of expressed recombinant L773P A is shown in SEQ ID NO:361, and the DNA sequence is shown in SEQ ID NO:362.

EXAMPLE 13

EPITOPES DERIVED FROM CLONE L773P POLYPEPTIDE

A series of peptides from the L773P amino acid sequence were synthesized and used in *in vitro* priming experiments to generate peptide-specific CD4 T cells. These peptides were 20-mers that overlapped by 15 amino acids and corresponded to amino acids 1-69 of the L773P protein. This region has been demonstrated to be tumor-specific.

- 25 Following three *in vitro* stimulations, CD4 T cell lines were identified that produced IFN γ in response to the stimulating peptide but not the control peptide. Some of these T cell lines demonstrated recognition of recombinant L773P and L773PA (tumor-sprcidic region) proteins.

To perform the experiments, a total of 11 20-mer peptides overlapping by 15 amino acids and derived from the N-terminal tumor-specific region of L773P corresponding to amino acids 1-69 of L773P were generated by standard procedures (Figure 1). Dendritic cells were derived from PBMC of a normal donor using GMCSF and IL-4 by standard protocol. Purified CD4 T cells were generated from the same donor as the dendritic cells by using MACS beads and negative selection of PBMCs. Dendritic cells were pulsed overnight with the individual 20-mer peptides at a concentration of 10 µg/ml. Pulsed dendritic cells were washed and plated at 1×10^4 /well of a 96-well U-bottom plates, and purified CD4 cells were added at 1×10^5 well. Cultures were supplemented with 10 ng/ml IL-6 and 5 ng/ml IL-12 and incubated at 37°C. Cultures were re-stimulated as above on a weekly basis using as APC dendritic cells generated and pulsed as above, supplemented with 5 ng/ml IL-7 and 10 µg/ml IL-2. Following 3 *in vitro* stimulation cycles, lines (each line corresponds to one well) were tested for cytokine production in response to the stimulating peptide vs. an irrelevant peptide.

A small number of individual CD4 T cell lines (9/528) demonstrated cytokine release (IFN γ) in response to the stimulating peptide but not to control peptide (Figure 3). The CD4 T cell lines that demonstrated specific activity were restimulated on the appropriate L773P peptide and reassayed using autologous dendritic cells pulsed with 10 µg/ml of the appropriate L773P peptide, an irrelevant control peptide, recombinant L773P protein (amino acids 2-364, made in *E. coli*), recombinant L773PA (amino acids 2-71, made in *E. coli*), and an appropriate control protein (L3E, made in *E. coli*). Three of the nine lines tested (1-3C, 1-6G, and 4-12B) recognized the appropriate L773P peptide as well as recombinant L773P and L773PA (Figure 2). Four of the lines tested (4-8A, 4-8E, 4-12D, and 4-12E) recognized the appropriate L773P peptide only. Two of the lines tested (5-6F and 9-3B) demonstrated non-specific activity.

The significant conclusion of this study is that the peptide sequences MWQPLFFKWLLSCCPGSSQI (amino acids 1-20, SEQ ID NO:363) and GSSQIAAAASTQPEDDINTQ (amino acids 16-35, SEQ ID NO: 365) may represent naturally processed epitopes of L773P, which are capable of stimulating human class II MHC-restricted CD4 T cell responses.

On the basis of these results, other epitopes within the scope of the invention include epitopes restricted by other class II MHC; molecules. In addition, variants of the peptide can be produced wherein one or more amino acids are altered such that there is no effect on the ability of the peptides to bind to MHC molecules, no effect on their ability to
 5 elicit T cell responses, and no effect on the ability of the elicited T cells to recognize recombinant protein.

The identification of these epitopes from L773P provides strong evidence that this antigen could be used as a component of a cancer vaccine for eliciting T cell responses in lung cancer patients for the treatment of their disease. The peptides could also
 10 be used for clinical monitoring of L773P vaccine-treated patients. The peptides could be used directly as a vaccine for lung cancer patients with an L773P-expressing lung tumor.

EXAMPLE 14

SURFACE EXPRESSION OF L762P AND ANTIBODY EPITOPES THEREOF

15 Rabbits were immunized with full-length Histidine-tagged L762 protein generated in E. coli. Sera was isolated from rabbits and screened for specific recognition of L762P in ELISA assays. One polyclonal serum, 2692L was identified that specifically recognized recombinant L762P protein. The 2692L anti-L762P polyclonal antibodies were purified from the serum by affinity purification using L762P affinity columns. Although
 20 L762P is expressed in a subset of primary lung tumor samples, expression appears to be lost in established lung tumor cell lines. Therefore, to characterize surface expression of L762P, a retrovirus construct that expresses L762P was used to transduce primary human fibroblasts as well as 3 lung tumor cell lines (522-23, HTB, and 343T). Transduced lines were selected and expanded to examine L762P surface expression by FACS analysis. For
 25 this analysis, non-transduced and transduced cells were harvested using cell dissociation medium, and incubated with 10-50 micrograms/ml of either affinity purified anti-L762P or irrelevant anti-P703P sera. Following a 30 minute incubation on ice, cells were washed and incubated with a secondary, FITC conjugated anti rabbit IgG antibody as above. Cells were washed, resuspended in buffer with Propidium Iodide (PI) and examined by FACS using an

Excalibur fluorescence activated cell sorter. For FACS analysis, PI-positive (i.e. dead/permeabilized cells) were excluded. The polyclonal anti-L762P sera specifically recognized and bound to the surface of L762P-transduced cells but not the non-transduced counterparts. These results demonstrate that L762P is localized to the cell surface of both fibroblasts as well as lung tumor cells.

To identify the peptide epitopes recognized by 2692L, an epitope mapping approach was pursued. A series of overlapping 19-21 mers (5 amino acid overlap) was synthesized that spanned C terminal 1/2 of L762P (amino acids 481-894). In an initial experiment peptides were tested in pools. Specific reactivity with the L762P antiserum was observed with pools A, B, C, and E. To identify the specific peptides recognized by the antiserum, flat bottom 96 well microtiter plates were coated with individual peptides at 10 microgram/ml for 2 hours at 37 C. Wells were then aspirated and blocked with phosphate buffered saline containing 5% (w/v) milk for 2 hours at 37 C, and subsequently washed in PBS containing 0.1% Tween 20 (PBST). Purified rabbit anti-L762P serum 2692L was added at 200 or 20 ng/well to triplicate wells in PBST and incubated overnight at room temperature. This was followed by washing 6 times with PBST and subsequently incubating with HRP-conjugated donkey anti rabbit IgG (H+L)Affinipure F(ab') fragment at 1:2,000 for 60 minutes. Plates were then washed, and incubated in Tetramethyl benzidine substrate. Reactions were stopped by the addition of 1N sulfuric acid and plates were read at 450/570 nm using an ELISA plate reader.

The resulting data, presented in Table 1 below, demonstrates that the L762P antisera recognized at least 6 distinct peptide epitopes from the 3' half of L762P.

Peptide (starting amino acid of L762P)	pool	ELISA activity (OD 450-570)	
		200 ng polyclonal serum	20 ng polyclonal serum
A (481)	A	1.76	1.0
B (495)	A	0.14	.06
C (511)	E	0.47	0.18
D (526)	E	0.11	0.09
E (541)	A	0.11	0.04
F (556)	A	0.04	0.02
G (571)	A	0.06	0.02
H (586)	B	0.1	0.03
I (601)	B	0.25	0.06
J (616)	B	0.1	0.03
K (631)	E	0.1	0.08
L (646)	B	0.28	0.12
M (661)	B	0.14	0.03
N (676)	C	0.12	0.1
O (691)	C	1.1	0.23
P (706)	C	0.1	0.03
Q (721)	C	0.11	0.05
R (736)	E	0.12	0.04
S (751)	C	0.15	0.06
U (781)	D	0.12	0.06
V (795)	F	0.07	0.05
X (826)	D	0.1	0.03
Y (841)	D	0.17	0.07
Z (856)	D	0.16	0.08
AA (871)	F	0.17	0.05
BB (874)	F	0.14	0.11
No peptide		0.15	0.045

Individual peptides were identified from each of the pools, and additionally a weak reactivity was identified with peptide BB from pool F. The relevant peptide

5 epitopes are summarized in the table below.

Peptide	Nucleotides of L762P	Amino acids of L762P	Sequence	pool	ELISA activity (OD 450-570)	
					200 ng	20 ng
A	1441-1500	481-500	SRISSTGDIQFQQHIQLEST	A	1.76	1.0
C	1531-1590	511-530	KNTVTVDNTVGNDTMFLVTW	E	0.47	0.18
I	1801-1860	601-620	AVPPATVEAFVERDSLHFPH	B	0.25	0.06
L	1936-1955	646-665	PETGDPVTLRLDDGAGADV	B	0.28	0.12
O	2071-2130	691-710	VNHSPSISTPAHSIPGSHAMIL	C	1.1	0.23
BB	2620-2679	874-893	LQSAVSNIQAPLFIPPNSD	F	0.14	0.11
None	-	-	-	-	0.15	0.05

From the foregoing it will be appreciated that, although specific
 5 embodiments of the invention have been described herein for purposes of illustration,
 various modifications may be made without deviating from the spirit and scope of the
 invention. Accordingly, the invention is not limited except as by the appended claims.

CLAIMS

What is claimed:

1. An isolated polypeptide, comprising at least an immunogenic portion of a lung tumor protein, or a variant thereof, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

(a) sequences recited in SEQ ID NO: 1-3, 6-8, 10-13, 15-27, 29, 30, 32, 34-49, 51, 52, 54, 55, 57-59, 61-69, 71, 73, 74, 77, 78, 80-82, 84, 86-96, 107-109, 111, 113, 125, 127, 128, 129, 131-133, 142, 144, 148-151, 153, 154, 157, 158, 160, 167, 168, 171, 173, 175, 179, 182, 184-186, 188-191, 193, 194, 198-207, 209, 210, 213, 214, 217, 220-224, 253, 254-258, 260, 262-264, 270, 272, 275, 276, 279-281, 286, 287, 291, 293, 295, 296, 300, 302, 308-310, 313, 315-317, 323, 345, 347 and 349;

(b) sequences that hybridize to a sequence recited in any one of SEQ ID NO: 1-3, 6-8, 10-13, 15-27, 29, 30, 32, 34-49, 51, 52, 54, 55, 57-59, 61-69, 71, 73, 74, 77, 78, 80-82, 84, 86-96, 107-109, 111, 113, 125, 127, 128, 129, 131-133, 142, 144, 148-151, 153, 154, 157, 158, 160, 167, 168, 171, 173, 175, 179, 182, 184-186, 188-191, 193, 194, 198-207, 209, 210, 213, 214, 217, 220-224, 253, 254-258, 260, 262-264, 270, 272, 275, 276, 279-281, 286, 287, 291, 293, 295, 296, 300, 302, 308-310, 313, 315-317, 323, 345, 347 and 349 under moderately stringent conditions; and

(c) complements of sequences of (a) or (b).

2. An isolated polypeptide according to claim 1, wherein the polypeptide comprises an amino acid sequence that is encoded by a polynucleotide sequence recited in any one of SEQ ID NO: 1-3, 6-8, 10-13, 15-27, 29, 30, 32, 34-49, 51, 52, 54, 55, 57-59, 61-69, 71, 73, 74, 77, 78, 80-82, 84, 86-96, 107-109, 111, 113, 125, 127, 128, 129, 131-133, 142, 144, 148-151, 153, 154, 157, 158, 160, 167, 168, 171, 173, 175, 179, 182, 184-186, 188-191, 193, 194, 198-207, 209, 210, 213, 214, 217, 220-224, 253, 254-258, 260, 262-264, 270, 272, 275, 276, 279-281, 286, 287, 291, 293, 295, 296, 300, 302, 308-310, 313, 315-317, 323, 345, 347 and 349 or a complement of any of the foregoing polynucleotide sequences.

3. An isolated polypeptide comprising a sequence recited in any one of SEQ ID NO: 110, 112, 114, 152, 155, 156, 159, 161, 165, 166, 169, 170, 172, 174, 176, 226-252, 346, 348 and 350.

4. An isolated polynucleotide encoding at least 15 amino acid residues of a lung tumor protein, or a variant thereof that differs in one or more substitutions, deletions, additions and/or insertions such that the ability of the variant to react with antigen-specific antisera is not substantially diminished, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide comprising a sequence recited in any one of SEQ ID NO: 1-3, 6-8, 10-13, 15-27, 29, 30, 32, 34-49, 51, 52, 54, 55, 57-59, 61-69, 71, 73, 74, 77, 78, 80-82, 84, 86-96, 107-109, 111, 113, 125, 127, 128, 129, 131-133, 142, 144, 148-151, 153, 154, 157, 158, 160, 167, 168, 171, 173, 175, 179, 182, 184-186, 188-191, 193, 194, 198-207, 209, 210, 213, 214, 217, 220-224, 253, 254-258, 260, 262-264, 270, 272, 275, 276, 279-281, 286, 287, 291, 293, 295, 296, 300, 302, 308-310, 313, 315-317, 323, 345, 347 and 349 or a complement of any of the foregoing sequences.

5. An isolated polynucleotide encoding a lung tumor protein, or a variant thereof, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide comprising a sequence recited in any one of SEQ ID NO: 1-3, 6-8, 10-13, 15-27, 29, 30, 32, 34-49, 51, 52, 54, 55, 57-59, 61-69, 71, 73, 74, 77, 78, 80-82, 84, 86-96, 107-109, 111, 113, 125, 127, 128, 129, 131-133, 142, 144, 148-151, 153, 154, 157, 158, 160, 167, 168, 171, 173, 175, 179, 182, 184-186, 188-191, 193, 194, 198-207, 209, 210, 213, 214, 217, 220-224, 253, 254-258, 260, 262-264, 270, 272, 275, 276, 279-281, 286, 287, 291, 293, 295, 296, 300, 302, 308-310, 313, 315-317, 323, 345, 347 and 349 or a complement of any of the foregoing sequences.

6. An isolated polynucleotide, comprising a sequence recited in any one of SEQ ID NO: 1-3, 6-8, 10-13, 15-27, 29, 30, 32, 34-49, 51, 52, 54, 55, 57-59, 61-69, 71, 73, 74, 77, 78, 80-82, 84, 86-96, 107-109, 111, 113, 125, 127, 128, 129, 131-133, 142, 144, 148-151,

153, 154, 157, 158, 160, 167, 168, 171, 173, 175, 179, 182, 184-186, 188-191, 193, 194, 198-207, 209, 210, 213, 214, 217, 220-224, 253, 254-258, 260, 262-264, 270, 272, 275, 276, 279-281, 286, 287, 291, 293, 295, 296, 300, 302, 308-310, 313, 315-317, 323, 345, 347 and 349.

7. An isolated polynucleotide, comprising a sequence that hybridizes to a sequence recited in any one of SEQ ID NO: 1-3, 6-8, 10-13, 15-27, 29, 30, 32, 34-49, 51, 52, 54, 55, 57-59, 61-69, 71, 73, 74, 77, 78, 80-82, 84, 86-96, 107-109, 111, 113, 125, 127, 128, 129, 131-133, 142, 144, 148-151, 153, 154, 157, 158, 160, 167, 168, 171, 173, 175, 179, 182, 184-186, 188-191, 193, 194, 198-207, 209, 210, 213, 214, 217, 220-224, 253, 254-258, 260, 262-264, 270, 272, 275, 276, 279-281, 286, 287, 291, 293, 295, 296, 300, 302, 308-310, 313, 315-317, 323, 345, 347 and 349 under moderately stringent conditions.

8. An isolated polynucleotide complementary to a polynucleotide according to any one of claims 4-7.

9. An expression vector, comprising a polynucleotide according to any one of claims claim 4-8.

10. A host cell transformed or transfected with an expression vector according to claim 9.

11. An isolated antibody, or antigen-binding fragment thereof, that specifically binds to a lung tumor protein that comprises an amino acid sequence that is encoded by a polynucleotide sequence recited in any one of SEQ ID NO: 1-3, 6-8, 10-13, 15-27, 29, 30, 32, 34-49, 51, 52, 54, 55, 57-59, 61-69, 71, 73, 74, 77, 78, 80-82, 84, 86-96, 107-109, 111, 113, 125, 127, 128, 129, 131-133, 142, 144, 148-151, 153, 154, 157, 158, 160, 167, 168, 171, 173, 175, 179, 182, 184-186, 188-191, 193, 194, 198-207, 209, 210, 213, 214, 217, 220-224, 253, 254-258, 260, 262-264, 270, 272, 275, 276, 279-281, 286, 287, 291, 293, 295, 296, 300, 302,

308-310, 313, 315-317, 323, 345, 347 and 349 or a complement of any of the foregoing polynucleotide sequences.

12. A fusion protein, comprising at least one polypeptide according to claim 1.

13. A fusion protein according to claim 12, wherein the fusion protein comprises an expression enhancer that increases expression of the fusion protein in a host cell transfected with a polynucleotide encoding the fusion protein.

14. A fusion protein according to claim 12, wherein the fusion protein comprises a T helper epitope that is not present within the polypeptide of claim 1.

15. A fusion protein according to claim 12, wherein the fusion protein comprises an affinity tag.

16. An isolated polynucleotide encoding a fusion protein according to claim 12.

17. A pharmaceutical composition, comprising a physiologically acceptable carrier and at least one component selected from the group consisting of:

- (a) a polypeptide according to claim 1;
- (b) a polynucleotide according to claim 4;
- (c) an antibody according to claim 11;
- (d) a fusion protein according to claim 12; and
- (e) a polynucleotide according to claim 16.

18. A vaccine comprising an immunostimulant and at least one component selected from the group consisting of:

- (a) a polypeptide according to claim 1;

- (b) a polynucleotide according to claim 4;
- (c) an antibody according to claim 11;
- (d) a fusion protein according to claim 12; and
- (e) a polynucleotide according to claim 16.

19. A vaccine according to claim 18, wherein the immunostimulant is an adjuvant.

20. A vaccine according to any claim 18, wherein the immunostimulant induces a predominantly Type I response.

21. A method for inhibiting the development of a cancer in a patient, comprising administering to a patient an effective amount of a pharmaceutical composition according to claim 17.

22. A method for inhibiting the development of a cancer in a patient, comprising administering to a patient an effective amount of a vaccine according to claim 18.

23. A pharmaceutical composition comprising an antigen-presenting cell that expresses a polypeptide according to claim 1, in combination with a pharmaceutically acceptable carrier or excipient.

24. A pharmaceutical composition according to claim 23, wherein the antigen presenting cell is a dendritic cell or a macrophage.

25. A vaccine comprising an antigen-presenting cell that expresses a polypeptide comprising at least an immunogenic portion of a lung tumor protein, or a variant thereof, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

- (a) sequences recited in SEQ ID NO: 1-109, 111, 113, 115-151, 153, 154, 157, 158, 160, 162-164, 167, 168, 171, 173, 175, 177-224, 255-337, 345, 347, 349 and 368;
- (b) sequences that hybridize to a sequence recited in any one of SEQ ID NO: 1-109, 111, 113, 115-151, 153, 154, 157, 158, 160, 162-164, 167, 168, 171, 173, 175, 177-224, 255-337, 345, 347, 349 and 368 under moderately stringent conditions; and
- (c) complements of sequences of (i) or (ii);
in combination with an immunostimulant.

26. A vaccine according to claim 25, wherein the immunostimulant is an adjuvant.

27. A vaccine according to claim 25, wherein the immunostimulant induces a predominantly Type I response.

28. A vaccine according to claim 25, wherein the antigen-presenting cell is a dendritic cell.

29. A method for inhibiting the development of a cancer in a patient, comprising administering to a patient an effective amount of an antigen-presenting cell that expresses a polypeptide comprising at least an immunogenic portion of a lung tumor protein, or a variant thereof, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

- (a) sequences recited in SEQ ID NO: 1-109, 111, 113, 115-151, 153, 154, 157, 158, 160, 162-164, 167, 168, 171, 173, 175, 177-224, 255-337, 345, 347, 349 and 368;

(b) sequences that hybridize to a sequence recited in any one of SEQ ID NO: 1-109, 111, 113, 115-151, 153, 154, 157, 158, 160, 162-164, 167, 168, 171, 173, 175, 177-224, 255-337, 345, 347, 349 and 368 under moderately stringent conditions; and

(c) complements of sequences of (i) or (ii) encoded by a polynucleotide recited in any one of SEQ ID NO: 1-109, 111, 113, 115-151, 153, 154, 157, 158, 160, 162-164, 167, 168, 171, 173, 175, 177-224, 255-337, 345, 347, 349 and 368;

and thereby inhibiting the development of a cancer in the patient.

30. A method according to claim 29, wherein the antigen-presenting cell is a dendritic cell.

31. A method according to any one of claims 21, 22 and 29, wherein the cancer is lung cancer.

32. A method for removing tumor cells from a biological sample, comprising contacting a biological sample with T cells that specifically react with a lung tumor protein, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

(i) polynucleotides recited in any one of SEQ ID NO: 1-109, 111, 113, 115-151, 153, 154, 157, 158, 160, 162-164, 167, 168, 171, 173, 175, 177-224, 255-337, 345, 347, 349 and 368; and

(ii) complements of the foregoing polynucleotides;

wherein the step of contacting is performed under conditions and for a time sufficient to permit the removal of cells expressing the antigen from the sample.

33. A method according to claim 32, wherein the biological sample is blood or a fraction thereof.

34. A method for inhibiting the development of a cancer in a patient, comprising administering to a patient a biological sample treated according to the method of claim 32.

35. A method for stimulating and/or expanding T cells specific for a lung tumor protein, comprising contacting T cells with at least one component selected from the group consisting of:

(a) polypeptides comprising at least an immunogenic portion of a lung tumor protein, or a variant thereof, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

(i) sequences recited in SEQ ID NO: 1-109, 111, 113, 115-151, 153, 154, 157, 158, 160, 162-164, 167, 168, 171, 173, 175, 177-224, 255-337, 345, 347, 349 and 368;

(ii) sequences that hybridize to a sequence recited in any one of SEQ ID NO: 1-109, 111, 113, 115-151, 153, 154, 157, 158, 160, 162-164, 167, 168, 171, 173, 175, 177-224, 255-337, 345, 347, 349 and 368 under moderately stringent conditions; and

(iii) complements of sequences of (i) or (ii);

(b) polynucleotides encoding a polypeptide of (a); and

(c) antigen presenting cells that express a polypeptide of (a);

under conditions and for a time sufficient to permit the stimulation and/or expansion of T cells.

36. An isolated T cell population, comprising T cells prepared according to the method of claim 35.

37. A method for inhibiting the development of a cancer in a patient, comprising administering to a patient an effective amount of a T cell population according to claim 36.

38. A method for inhibiting the development of a cancer in a patient, comprising the steps of:

(a) incubating CD4⁺ and/or CD8⁺ T cells isolated from a patient with at least one component selected from the group consisting of:

(i) polypeptides comprising at least an immunogenic portion of a lung tumor protein, or a variant thereof, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

(1) sequences recited in SEQ ID NO: 1-109, 111, 113, 115-151, 153, 154, 157, 158, 160, 162-164, 167, 168, 171, 173, 175, 177-224, 255-337, 345, 347, 349 and 368;

(2) sequences that hybridize to a sequence recited in any one of SEQ ID NO: 1-109, 111, 113, 115-151, 153, 154, 157, 158, 160, 162-164, 167, 168, 171, 173, 175, 177-224, 255-337, 345, 347, 349 and 368 under moderately stringent conditions; and

(3) complements of sequences of (1) or (2);

(ii) polynucleotides encoding a polypeptide of (i); and

(iii) antigen presenting cells that expresses a polypeptide of (i);

such that T cells proliferate; and

(b) administering to the patient an effective amount of the proliferated T cells, and thereby inhibiting the development of a cancer in the patient.

39. A method for inhibiting the development of a cancer in a patient, comprising the steps of:

(a) incubating CD4⁺ and/or CD8⁺ T cells isolated from a patient with at least one component selected from the group consisting of:

(i) polypeptides comprising at least an immunogenic portion of a lung tumor protein, or a variant thereof, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence selected from the group consisting of:

(1) sequences recited in SEQ ID NO: 1-109, 111, 113, 115-151, 153, 154, 157, 158, 160, 162-164, 167, 168, 171, 173, 175, 177-224, 255-337, 345, 347, 349 and 368;

(2) sequences that hybridize to a sequence recited in any one of SEQ ID NO: 1-109, 111, 113, 115-151, 153, 154, 157, 158, 160, 162-164, 167, 168, 171, 173, 175, 177-224, 255-337, 345, 347, 349 and 368 under moderately stringent conditions; and

(3) complements of sequences of (1) or (2);

(ii) polynucleotides encoding a polypeptide of (i); and

(iii) antigen presenting cells that express a polypeptide of (i);

such that T cells proliferate;

(b) cloning at least one proliferated cell to provide cloned T cells; and

(c) administering to the patient an effective amount of the cloned T cells, and thereby inhibiting the development of a cancer in the patient.

40. A method for determining the presence or absence of a cancer in a patient, comprising the steps of:

(a) contacting a biological sample obtained from a patient with a binding agent that binds to a lung tumor protein, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence recited in any one of SEQ ID NO: 1-109, 111, 113, 115-151, 153, 154, 157, 158, 160, 162-164, 167, 168, 171, 173, 175, 177-224, 255-337, 345, 347, 349 and 368 or a complement of any of the foregoing polynucleotide sequences;

(b) detecting in the sample an amount of polypeptide that binds to the binding agent; and

(c) comparing the amount of polypeptide to a predetermined cut-off value, and therefrom determining the presence or absence of a cancer in the patient.

41. A method according to claim 40, wherein the binding agent is an antibody.

42. A method according to claim 41, wherein the antibody is a monoclonal antibody.

43. A method according to claim 40, wherein the cancer is lung cancer.

44. A method for monitoring the progression of a cancer in a patient, comprising the steps of:

(a) contacting a biological sample obtained from a patient at a first point in time with a binding agent that binds to a lung tumor protein, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence recited in any one of SEQ ID NO: 1-109, 111, 113, 115-151, 153, 154, 157, 158, 160, 162-164, 167, 168, 171, 173, 175, 177-224, 255-337, 345, 347, 349 and 368 or a complement of any of the foregoing polynucleotide sequences;

(b) detecting in the sample an amount of polypeptide that binds to the binding agent;

(c) repeating steps (a) and (b) using a biological sample obtained from the patient at a subsequent point in time; and

(d) comparing the amount of polypeptide detected in step (c) to the amount detected in step (b) and therefrom monitoring the progression of the cancer in the patient.

45. A method according to claim 44, wherein the binding agent is an antibody.

46. A method according to claim 45, wherein the antibody is a monoclonal antibody.

47. A method according to claim 44, wherein the cancer is a lung cancer.

48. A method for determining the presence or absence of a cancer in a patient, comprising the steps of:

(a) contacting a biological sample obtained from a patient with an oligonucleotide that hybridizes to a polynucleotide that encodes a lung tumor protein, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence recited in any one of SEQ ID NO: 1-109, 111, 113, 115-151, 153, 154, 157, 158, 160, 162-164, 167, 168, 171, 173, 175, 177-224, 255-337, 345, 347, 349 and 368 or a complement of any of the foregoing polynucleotide sequences;

(b) detecting in the sample an amount of a polynucleotide that hybridizes to the oligonucleotide; and

(c) comparing the amount of polynucleotide that hybridizes to the oligonucleotide to a predetermined cut-off value, and therefrom determining the presence or absence of a cancer in the patient.

49. A method according to claim 48, wherein the amount of polynucleotide that hybridizes to the oligonucleotide is determined using a polymerase chain reaction.

50. A method according to claim 48, wherein the amount of polynucleotide that hybridizes to the oligonucleotide is determined using a hybridization assay.

51. A method for monitoring the progression of a cancer in a patient, comprising the steps of:

(a) contacting a biological sample obtained from a patient with an oligonucleotide that hybridizes to a polynucleotide that encodes a lung tumor protein, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence recited in any one of SEQ ID NO: 1-109, 111, 113, 115-151, 153, 154, 157, 158, 160, 162-164, 167, 168, 171, 173, 175, 177-224, 255-337, 345, 347, 349 and 368 or a complement of any of the foregoing polynucleotide sequences;

(b) detecting in the sample an amount of a polynucleotide that hybridizes to the oligonucleotide;

(c) repeating steps (a) and (b) using a biological sample obtained from the patient at a subsequent point in time; and

(d) comparing the amount of polynucleotide detected in step (c) to the amount detected in step (b) and therefrom monitoring the progression of the cancer in the patient.

52. A method according to claim 51, wherein the amount of polynucleotide that hybridizes to the oligonucleotide is determined using a polymerase chain reaction.

53. A method according to claim 51, wherein the amount of polynucleotide that hybridizes to the oligonucleotide is determined using a hybridization assay.

54. A diagnostic kit, comprising:

- (a) one or more antibodies according to claim 11; and
- (b) a detection reagent comprising a reporter group.

55. A kit according to claim 54, wherein the antibodies are immobilized on a solid support.

56. A kit according to claim 54, wherein the detection reagent comprises an anti-immunoglobulin, protein G, protein A or lectin.

57. A kit according to claim 54, wherein the reporter group is selected from the group consisting of radioisotopes, fluorescent groups, luminescent groups, enzymes, biotin and dye particles.

58. An oligonucleotide comprising 10 to 40 contiguous nucleotides that hybridize under moderately stringent conditions to a polynucleotide that encodes a lung tumor

protein, wherein the tumor protein comprises an amino acid sequence that is encoded by a polynucleotide sequence recited in any one of SEQ ID NO: 1-3, 6-8, 10-13, 15-27, 29, 30, 32, 34-49, 51, 52, 54, 55, 57-59, 61-69, 71, 73, 74, 77, 78, 80-82, 84, 86-96, 107-109, 111, 113, 125, 127, 128, 129, 131-133, 142, 144, 148-151, 153, 154, 157, 158, 160, 167, 168, 171, 173, 175, 179, 182, 184-186, 188-191, 193, 194, 198-207, 209, 210, 213, 214, 217, 220-224, 253, 254-258, 260, 262-264, 270, 272, 275, 276, 279-281, 286, 287, 291, 293, 295, 296, 300, 302, 308-310, 313, 315-317, 323, 345, 347 and 349 or a complement of any of the foregoing polynucleotides.

59. A oligonucleotide according to claim 58, wherein the oligonucleotide comprises 10-40 contiguous nucleotides recited in any one of SEQ ID NO: 1-3, 6-8, 10-13, 15-27, 29, 30, 32, 34-49, 51, 52, 54, 55, 57-59, 61-69, 71, 73, 74, 77, 78, 80-82, 84, 86-96, 107-109, 111, 113, 125, 127, 128, 129, 131-133, 142, 144, 148-151, 153, 154, 157, 158, 160, 167, 168, 171, 173, 175, 179, 182, 184-186, 188-191, 193, 194, 198-207, 209, 210, 213, 214, 217, 220-224, 253, 254-258, 260, 262-264, 270, 272, 275, 276, 279-281, 286, 287, 291, 293, 295, 296, 300, 302, 308-310, 313, 315-317, 323, 345, 347 and 349.

60. A diagnostic kit, comprising:

- (a) an oligonucleotide according to claim 59; and
- (b) a diagnostic reagent for use in a polymerase chain reaction or hybridization assay.

COMPOSITIONS AND METHODS FOR THE THERAPY
AND DIAGNOSIS OF LUNG CANCER

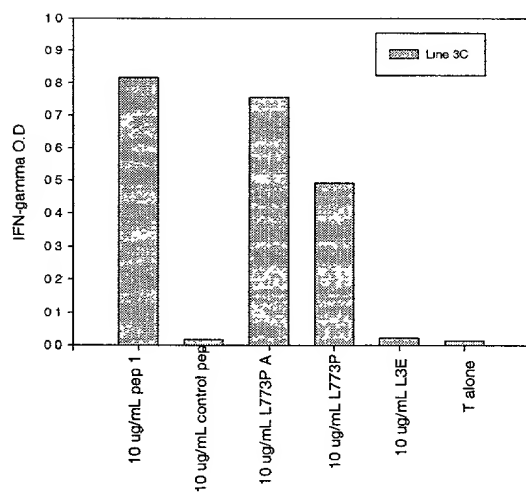
ABSTRACT OF THE DISCLOSURE

Compositions and methods for the therapy and diagnosis of cancer, such as lung cancer, are disclosed. Compositions may comprise one or more lung tumor proteins, immunogenic portions thereof, or polynucleotides that encode such portions. Alternatively, a therapeutic composition may comprise an antigen presenting cell that expresses a lung tumor protein, or a T cell that is specific for cells expressing such a protein. Such compositions may be used, for example, for the prevention and treatment of diseases such as lung cancer. Diagnostic methods based on detecting a lung tumor protein, or mRNA encoding such a protein, in a sample are also provided.

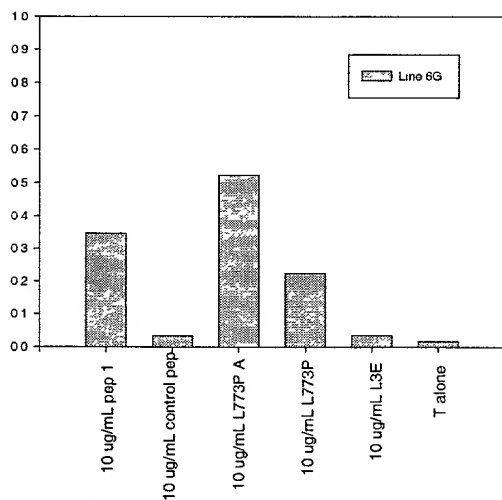
WPN\210121\corixa\455c13\455c13-ap.doc

L773P Peptides	
MWQPLFFKWLLSCCPGSSQI	1-20
FFKWLLSCCPGSSQIAAAAS	6-25
LSCCPGSSQIAAAASTQPED	11-30
GSSQIAAAASTQPEDDINTQ	16-35
AAASTQPEDDINTQRKKSQ	21-40
TQPEDDINTQRKKSQEKMRE	26-45
DINTQRKKSQEKMREVTDSP	31-50
RKKSQEKMREVTDSPGRPRE	36-55
EKMREVTDSPGRPRELTIPQ	41-60
VTDSPGRPRELTIPQTSSHG	46-65
GRPRELTIPQTSSHGANRF	51-69

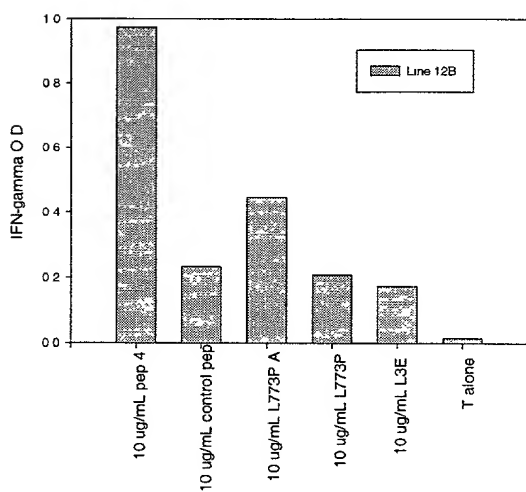
Fig. 1



2A



2B



2C

Fig. 2

100%
 80%
 60%
 40%
 20%
 0%
 0 10 20 30 40 50 60 70 80 90 100

D45 L773 CD4 Assay					
IFN-gamma SI					
Peptide 1					
SI	SI	SI	SI	SI	SI
07	01	07	12	05	12
09	16	20	14	08	13
08	09	6.4	07	11	07
12	12	08	05	19	08
08	06	08	08	11	20
08	09	13	11	13	07
12	12	09	05	10	4.2
14	12	06	06	06	24
Peptide 2					
09	11	08	05	2.7	11
24	10	14	08	09	13
07	09	09	14	10	16
08	10	11	12	09	84
06	21	09	15	09	15
10	13	11	16	10	12
09	10	10	10	11	08
21	08	05	08	10	15
Peptide 3					
10	09	15	13	11	09
11	11	10	08	12	08
13	11	09	09	20	11
13	04	13	14	09	11
15	06	13	07	11	09
08	15	13	06	13	10
07	11	16	09	23	05
10	25	09	24	09	09
Peptide 4					
09	13.6	06	08	09	10
09	39	12	09	15	13.7
10	09	09	07	05	11
09	08	07	11	09	12.3
08	2.8	09	10	12	4.3
11	08	10	12	07	12
11	10	11	10	10	12
09	07	05	07	11	11
Peptide 5					
12	07	10	07	21	12
10	09	09	12	07	08
09	09	08	14	13	15
09	17	11	13	15	15
07	08	09	05	14	13
06	10	11	10	08	31
10	08	11	10	12	06
08	11	07	08	08	17
Peptide 6					
08	08	08	10	10	15
11	08	07	14	10	11
11	08	14	07	07	25
08	11	09	08	11	12
07	15	11	08	11	07
13	10	09	27	14	11
11	10	08	11	16	18
09	13	19	10	08	12
Peptide 7					
08	22	09	06	10	13
12	12	11	08	16	09
13	14	12	14	08	11
22	09	12	09	13	08
13	10	07	17	06	07
10	17	18	16	10	12
07	08	12	10	15	12
08	16	06	07	15	21
Peptide 8					
09	08	10	07	10	10
07	10	11	09	11	10
16	10	08	12	16	12
09	10	11	13	11	09
11	10	14	10	09	09
25	11	09	08	11	10
06	17	11	06	14	11
09	08	12	10	15	10
Peptide 9					
10	08	49	07	08	13
09	09	24.4	12	13	13
10	11	26.3	12	18	12
06	07	28.1	11	08	10
11	10	13	07	12	11
07	09	40.2	09	13	12
10	08	44.0	08	11	09
09	08	51	15	13	16
Peptide 10					
10	10	12	11	09	18
04	09	12	09	05	11
08	07	08	10	10	13
10	10	06	11	11	15
10	12	11	13	08	11
11	10	11	06	13	13
12	08	10	13	15	10
07	08	12	15	13	12
Peptide 11					
11	12	12	12	16	12
09	09	10	12	11	13
11	07	14	19	14	16
11	09	15	10	10	15
12	09	17	09	09	14
12	11	11	15	14	11
11	08	13	10	14	18
11	07	09	13	22	21

Fig. 3

SEQUENCE LISTING

<110> Wang, Tongtong
 Fan, Liqun
 Kalos, Michael D.
 Bangur, Chaitanya S.
 Hosken, Nancy
 Fanger, Gary R.
 Li, Samuel X.
 Wang, Aijun
 Skeiky, Yasir A.W.
 Henderson, Robert A.
 McNeill, Patricia D.

<120> COMPOSITIONS AND METHODS FOR THE THERAPY
 AND DIAGNOSIS OF LUNG CANCER

<130> 210121.455C13

<140> US

<141> 2000-10-09

<160> 381

<170> FastSEQ for Windows Version 3.0

<210> 1

<211> 315

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(315)

<223> n = A,T,C or G

<400> 1

gcagagacag actggtgggtt gaacctggag gtgccaaaaa agccagctgc gggcccagga	60
cagctgccgt gagactcccg atgtcacagg cagtctgtgt ggttacagcg cccctcagtg	120
ttcatctcca gcagagacaa cggaggaggc tcccaccagg acggttctca ttatttatat	180
gttaatatgt ttgtaaactc atgtacagtt ttttttgggg gggaagcaat gggaangta	240
naaattacaa atagaatcat ttgctgtaat ccttaaattg caaacggtca ggccacgtga	300
aaaaaaaaaa aaaaaa	315

<210> 2

<211> 380

<212> DNA

<213> Homo sapien

<400> 2

atttaggctt aagatthttgt ttacccttgt tactaaggag caaattagta ttaaagtata	60
--	----

```

atatatataa acaaatacaa aaagttttga gtgggttcagc ttttttattt tttttaatgg      120
cataactttt aacaacactg ctctgtaatg gggttgaactg tggactcag actgagataa      180
ctgaaatgag tggatgtata gtgttattgc ataattatcc cactatgaag caaagggact      240
ggataaattc ccagtctaga ttattagcct ttgttaacca tcaagcacct agaagaagaa      300
ttattggaat ttttgcctc tgtaactggc actttggggg gtgacttata ttttgccttt      360
gtaaaaaaaa aaaaaaaaaa                                     380

```

```

<210> 3
<211> 346
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(346)
<223> n = A,T,C or G

```

```

<400> 3
ttgtaagtat acaatttttag aaaggattaa atgttattga tcattttact gaatactgca      60
catcctcacc atacaccatc cactttccaa taacatttaa tcctttctaa aattgtaagt      120
atacaattgt actttctttg gattttcata acaaatacac catagactgt taattttatt      180
gaagtttcct taatggaatg agtcattttt gtcttggtct tttgaggcta cctttgcttt      240
gacttccaac aatttgatca tatagtgttg agctgtggaa atctttaagt ttattctata      300
gcaataattt ctattnnnag annccnggnn naaaannann annaaa                          346

```

```

<210> 4
<211> 372
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(372)
<223> n = A,T,C or G

```

```

<400> 4
actagtctca ttactccaga attatgctct tgtacctgtg tggctggggt tcttagtcgt      60
tggtttggtt tggttttttg aactggatat tagggtggtt cacagttcta atgtaagcac      120
tctcttctcc aagttgtgct ttgtggggac aatcattctt tgaacattag agaggaaggc      180
agttcaagct gttgaaaaga ctattgctta tttttgtttt taaagacctt cttgacgtca      240
tgtggacagt gcacgtgcct tacgctacat cttgttttct aggaagaagg ggatgonggg      300
aaggantggg tgctttgtga tggataaaac gnctaaataa cacaccttta cattttgaaa      360
aaaacaaaac aa                                             372

```

```

<210> 5
<211> 698
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(698)
<223> n = A,T,C or G

```

120 180 240 300 360 380
 60 120 180 240 300 346
 4 372
 60 120 180 240 300 360 372

<400> 5

actagtanga	tagaaacact	gtgtcccgag	agtaaggaga	gaagctacta	ttgattagag	60
cctaaccag	gttaactgca	agaagaggcg	ggatactttc	agctttccat	gtaactgtat	120
gcataaagcc	aatgtagtcc	agttttctaag	atcatgttcc	aagctaactg	aatcccactt	180
caatacacac	tcatgaactc	ctgatggaac	aataacaggc	ccaagcctgt	ggtagatgt	240
gcacacttgc	tagactcaga	aaaaatacta	ctctcataaa	tgggtgggag	tattttgggt	300
gacaacctac	tttgcttggc	tgagtgaagg	aatgatattc	atatnttcat	ttattccatg	360
gacatttagt	tagtgctttt	tatataccag	gcagtatgct	gagtgcact	cttgtgtata	420
tntccaaatn	ttngtncngt	cgctgcacat	atctgaaatc	ctatattaag	antttcccaa	480
natgangtcc	ctggtttttc	cacgccactt	gatcngtcaa	ngatctcacc	tctgtntgtc	540
ctaaaaccnt	ctnctnnang	gttagaengg	acctctcttc	tcccttcccg	aanaatnaag	600
tgtgngaaga	nancncncn	ccccctncn	tncnccctng	cengctnnnc	cncntgtngg	660
gggngccgcc	ccgcggggg	gacccccccn	ttttcccc			698

<210> 6

<211> 740

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(740)

<223> n = A,T,C or G

<400> 6

actagtcaaa	aatgctaaaa	taatttggga	gaaaatattt	tttaagtagt	gttatagttt	60
catgtttatc	ttttattatg	tnttgtgaag	ttgtgtcttt	tcactaatta	cctatactat	120
gccaatattt	ccttatatct	atccataaca	tttatactac	atttgtaaga	gaatatgcac	180
gtgaaactta	acactttata	aggtaaaaat	gaggtttcca	agatttaata	atctgatcaa	240
gttcttggtta	tttccaaaata	gaatggactt	ggctgtgtta	ggggctaagg	gagaagaaga	300
agataagggt	aaaagttggt	aatgacccaa	cattctaaaa	gaaatgcaaa	aaaaaattta	360
ttttcaagcc	ttcgaactat	ttaaggaaa	caaaatcatt	tcctanatgc	atatcatttg	420
tgagantttc	tcantaatat	cctgaatcat	tcatttcagc	tnaggcttca	tgttgactcg	480
atatgtcatc	tagggaaagt	ctatttcatg	gtccaaacct	gttgccatag	ttggttnaggc	540
tttcctttaa	ntgtgaanta	tnacangaa	attttctctt	tnanagttct	tnatagggtt	600
aggggtgtgg	gaaaagcttc	taacaatctg	tagtgttncg	tgttatctgt	ncagaaccan	660
aatnacggat	cgnangaagg	actgggtcta	tttacangaa	cgaatnatct	ngttnnntgt	720
gtnnncaact	ccngggagcc					740

<210> 7

<211> 670

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(670)

<223> n = A,T,C or G

<400> 7

gctggggagc	tgcgcatggc	ggteccccgt	gcagccatgg	ggccctcggc	gttggggccag	60
agcggccccg	tcgcatggc	cccgtggtgc	tcagtgagca	gcggccccgc	gcgctacgtg	120
cttgggatgc	aggagctgtt	ccggggccac	agcaagaccg	cgagttcctg	gcgcacagcg	180
ccaaggtgca	ctcgggtggc	tgaggttgcg	acgggcgtcg	cctacctcgg	ggtcttcgac	240

```

aagacgccac gtcttcttgc tgganaanga ccgttggtca aagaaaacaa ttatcgggga      300
catggggata gtgtggacca ctttgttggc atccaagtaa tcctgacctt tttgttacgg      360
cgtctggaga taaaaccatt cgcctctggg atgtgaggac tacaaaatgc attgccactg      420
tgaacactaa agggggagaac attaatatct gctggantcc tgatgggcan accattgctg      480
tagcnacaag gatgatgtgg tgactttatt gatgccaaag aaccccgctt caaagcaaaa      540
aaacanttcc aanttcgaag tcacnnaaat ctcttggaac aatgaacatn aatatnttct      600
tcctgacaat ggnccctggg tgtntcacat cctcagctnc cccaaaactg aancctgtnc      660
natccacccc                                     670

```

```

<210> 8
<211> 689
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(689)
<223> n = A,T,C or G

```

```

<400> 8
actagtatct aggaatgaac agtaaaagag gagcagttgg ctacttgatt acaacagagt      60
aatgaagta ctggatttgg gaaaacctgg ttttattaga acatatggaa tgaaagccta      120
cacctagcat tgcctactta gccccctgaa ttaacagagc ccaattgaga caaacccttg      180
gcaacaggaa attcaaggga gaaaaagtaa gcaacttggg ctaggatgag ctgactccct      240
tagagcaaag ganagacagc cccattacc aaataccatt tttgcctggg gcttgtgcag      300
ctggcagtgt tcctgcccc gcatggcacc ttatngtttt gatagcaact tcgttgaatt      360
ttcaccaact tattacttga aattataata tagcctgtcc gtttgctgtg tccaggctgt      420
gatatatntt cctagtgggt tgacttttaa aataaatnag gtttantttt ctcccccn      480
cnntnctncc nntcnctcnn cnntcccccc cnetcngtcc tccnnnnntn gggggggccn      540
ccccncggg ggacccccct ttgggtccctt agtgagggtt natggcccct ggnnttatcc      600
nggcctann tttccccgtn nnaaatgntt cccctcccca ntccnccac ctcaanccgg      660
aagcctaagt ttntaccctg ggggtccccc                                     689

```

```

<210> 9
<211> 674
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(674)
<223> n = A,T,C or G

```

```

<400> 9
gtccactctc ctttgagtgt actgtcttac tgtgcactct gtttttcaac tttctagata      60
taaaaaatgc ttgtttctata gtggagtaag agctcacaca cccaaggcag caagataact      120
gaaaaaagcg aggcctttttt gccaccttgg taaaggccag ttactgcta tagaactgct      180
ataagcctga agggaagtag ctatgagact ttccattttt cttagtcttc ccaataggct      240
ccttcatgga aaaaggcttc ctgtaataat tttcacctaa tgaattagca gtgtgattat      300
ttctgaaata agagacaaat tgggcccgcag agtcttctct tgatttaaaa taaacaaccc      360
aaagttttgt ttggtcttca ccaaaggaca tactctaggg ggtatgttgt tgaagacatt      420
caaaaacatt agctgttctg tctttcaatt tcaagttatt ttggagactg cctccatgtg      480
agttaattac tttgctctgg aactagcatt attgtcatta tcatcacatt ctgtcatcat      540
catctgaata atattgtgga tttccccctc tgcttgcata ttcttttgac tcctctggga      600

```

```

anaaatgtca aaaaaaaagg tcgatctact cngcaaggnc catctaata ctgcgctgga 660
aggaccnct gccc 674

```

```

<210> 10
<211> 346
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(346)
<223> n = A,T,C or G

```

```

<400> 10
actagtctgc tgatagaaag cactatacat cctattgttt ctttctttcc aaaatcagcc 60
ttctgtctgt aacaaaaatg tactttatag agatggagga aaaggtctaa tactacatag 120
ccttaagtgt ttctgtcatt gttcaagtgt attttctgta acagaaacat atttggaatg 180
tttttctttt ccccttataa attgtaattc ctgaaatact gctgctttaa aaagtccac 240
tgtcagatta tattatctaa caattgaata ttgtaaata acttgtctta cctctcaata 300
aaagggtact tttctatttan nnagnngnnn gnnnnataaa anaaaa 346

```

```

<210> 11
<211> 602
<212> DNA
<213> Homo sapien

```

```

<400> 11
actagtaaaa agcagcattg ccaaataatc cctaattttc cactaaaaat ataatgaaat 60
gatgttaagc tttttgaaaa gtttaggtta aacctactgt tgtagatta atgtatttgt 120
tgcttccctt tatctggaat gtggcattag cttttttatt ttaaccctct ttaattctta 180
ttcaattcca tgacttaagg ttggagagct aaacactggg atttttggat aacagactga 240
cagttttgca taattataat cggcattgta catagaaagg atatggctac cttttgttaa 300
atctgcactt tctaaatatc aaaaaaggga aatgaagtta taaatcaatt tttgtataat 360
ctgtttgaaa catgagtttt atttgcttaa tattagggtt ttgccccctt tctgtaagtc 420
tcttgggatc ctgtgtagaa ctgttctcat taaacaccaa acagttaagt ccattctctg 480
gtactagcta caaattcggg ttcatattct acttaacaat ttaaataaac tgaaatatat 540
ctagatggtc tacttctgtt catataaaaa caaaacttga tttccaaaaa aaaaaaaaaa 600
aa 602

```

```

<210> 12
<211> 685
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(685)
<223> n = A,T,C or G

```

```

<400> 12
actagtctctg tgaaagtaca actgaaggca gaaagtgtta ggattttgca tctaattgttc 60
attatcatgg tattgatgga cctaagaaaa taaaaattag actaagcccc caaataagct 120
gcatgcattt gtaacatgat tagtagattt gaatatag atgtagtatn ttgggtatct 180
agggtgttta tcattatgta aaggaattaa agtaaaggac tttgtagttg tttttattaa 240

```

atatgcatat	agtagagtgc	aaaaatatag	caaaaatana	aactaaaggt	agaaaagcat	300
tttagatatg	cottaatnta	nnaactgtgc	caggtggccc	tcggaataga	tgccaggcag	360
agaccagtgc	ctgggtgggtg	cctccccctg	tctgcccccc	tgaagaactt	ccctcacgtg	420
angtagtgcc	ctcgtaggtg	tcacgtggan	tantggganc	aggccgnncn	gtanaaagaa	480
ancanngtga	nagtttncnc	gtngangcng	aactgtccct	gngccnnnac	gctcccanaa	540
cntntccaat	ngacaatcga	gtttccnnnc	tcnngnaacc	tngccgnnnn	cnngeccnnc	600
cantntgnta	accccgcgcc	cggatcgctc	tcnntcggtt	ctcncncnaa	ngggntttcn	660
cnnccgccgt	cncnnccccg	cnnc				685

<210> 13
 <211> 694
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(694)
 <223> n = A,T,C or G

<400> 13						
cactagtcac	tcattagcgt	tttcaatagg	gctcttaagt	ccagtagatt	acgggtagtc	60
agttgacgaa	gatctggttt	acaagaacta	attaaatgtt	tcattgcatt	tttgtaagaa	120
cagaataatt	ttataaaatg	tttgtagttt	ataattgccc	aaaataattt	aaagacactt	180
tttctctgtg	tgtgcaaatg	tgtgtttgtg	atccattttt	tttttttttt	taggacacct	240
gtttactagc	tagctttaca	atatgccaaa	aaaggatttc	tccttgaccc	catccgtggt	300
tcacctctt	ttccccccat	gctttttgcc	ctagtttata	acaaaggaat	gatgatgatt	360
taaaaagtag	ttctgtatct	tcagtatctt	ggtcttccag	aacctctg	ttgggaaggg	420
gatcattttt	tactggtcac	ttccctttgg	agtgtactac	tttaacagat	ggaaagaact	480
cattggccat	ggaaacagcc	gangtggttg	gagccagcag	tgcatggcac	cgtccggcat	540
ctggcntgat	tgggtctggc	gocgtcattg	tcagcacagt	gccatgggac	atggggaana	600
ctgactgcac	ngccaatgg	tttcatgaag	aatacngcat	ncnngtgat	cacgtnancc	660
angacgctat	gggggncana	gggccanttg	cttc			694

<210> 14
 <211> 679
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(679)
 <223> n = A,T,C or G

<400> 14						
cagccgcctg	catctgtatc	cagcgccang	tcccgcaggt	cccagctgcg	cgcgcccccc	60
agtcccgnc	ccgttcggcc	cangctnagt	tagnccctac	catnccggtc	aaaggangca	120
ccaagtgcac	caaataacctg	cngtncggat	ntaaattcat	cttctggctt	gccgggattg	180
ctgtccntgc	cattggacta	nggetccgat	ncgactctca	gaccanganc	atcttcganc	240
naganactaa	tnatnatnt	tccagcttct	acacaggagt	ctatattctg	atcggatccg	300
gnccectent	gatgctgggtg	ggcttccctga	gctgctgcgg	ggctgtgcaa	gagtcccant	360
gcatgctggg	actgttcttc	ggcttctctc	tgggtgatatn	cgccattgaa	atacctgcgg	420
ccatctgggg	atattccact	ncgatnatgt	gattaaggaa	ntccacggag	ttttacaagg	480
acacgtacaa	cnacctgaaa	accnnggatg	anccccaccg	ggaancnctg	aangccatcc	540
actatgcgtt	gaactgcaat	ggtttggctg	gggnccttga	acaatttaat	cncatacatc	600

tggccccann aaaggacntn ctcgannect tcnccgtgna attcngttct gatnccatca 660
cagaagtctc gaacaatcc 679

<210> 15
<211> 695
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(695)
<223> n = A,T,C or G

<400> 15
actagtggat aaaggccagg gatgetgctc aacctcctac catgtacagg gacgtctccc 60
cattacaact acccaatccg aagtgtcaac tgtgtcagga ctaanaaacc ctggttttga 120
ttaaaaaagg gcoctgaaaaa aggggagcca caaatctgtc tgcttcctca cnttantcnt 180
tggcaaatna gcattctgtc tcnttggctg cngcctcanc ncaaaaaanc ngaactcnat 240
cnggccagg aatacatctc ncaatnaacn aaattganca aggcnnctgg aaatgcnga 300
tgggattatc ntccgcttgt tgancctcta agtttcnttc ccttcattcn accctgccag 360
ccnagttctg ttagaaaaaat gcngaattc naacnccgg tttcntactc ngaatttaga 420
tctncanaaa ctctcctggcc acnattcnaa ttnangnca cgnacanatn ccttccatna 480
anencacccc acntttgana gccangacaa tgactgcntn aantgaaggc ntgaaggaan 540
aactttgaaa ggaaaaaaa ctttgtttcc ggcccttcc aacncttctg tgttnancac 600
tgcttctng naacctgga agccngnga cagtgttaca tgttgttcta nnaaacngac 660
ncttnaatnt cnatcttccc nanaacgatt ncncc 695

<210> 16
<211> 669
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(669)
<223> n = A,T,C or G

<400> 16
cgccgaagca gcagcgcagg ttgtccccgt ttccccctcc ccttcccttc tccggttgcc 60
ttccccggcc ccttacctc cacagtcctc gtccccccat gtcccagaaa caagaagaag 120
agaacctgc ggaggagacc ggagaggaga agcaggacac gcaggagaaa gaaggtattc 180
tgcttgagag agctgaagag gcaaagctaa aggccaaata cccaagccta ggacaaaagc 240
ctggaggctc cgacttctc atgaagagac tccagaaagg gcaaaagtac tttgactcng 300
gagactacaa catggccaaa gccaacatga agaataagca gctgccaagt gcangaccag 360
acaagaacct ggtgactggt gatcacatcc ccacccaca ggatctgcc agagaaagtc 420
ctcgtctgtc accagcaagc ttgcgggtgg ccaagttgaa tgatgctgcc ggggctctgc 480
canatctgag acgttccct ccctgcccc cccgggtcct gtgctggctc ctgcccttcc 540
tgcttttgca gccangggtc aggaagtggc ncnggtngtg gctggaaagc aaaaccttt 600
cctgttggtg tcccacccat ggagcccctg gggcgagccc angaacttga ncctttttgt 660
tntcttncc 669

<210> 17
<211> 697
<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(697)

<223> n = A,T,C or G

<400> 17

gcaagatatg	gacaactaag	tgagaaggta	atnctctact	gctctagntn	ctccnggcnn	60
gacgcgtga	ggagannnac	gctggcccan	ctgccggcca	cacacgggga	tcntggtnat	120
gectgcccan	gggancccca	ncnctcggan	cccatntcac	accggnnccn	tnccgcccacn	180
ncctggctcn	cncngcccng	nccagctenc	gncccccctcc	gcennnctcn	ttnnctcttc	240
cncnccctcc	ncnacnacct	cctaccencg	gctccctccc	cagccccccc	cgcgaancct	300
ccacnacncc	ntcnnencga	ancnecnetc	genctcngcc	ccngccccct	gccccccgcc	360
cncnacnneg	cgntcccccg	cgncgcngc	ctnccccct	cccacnacag	ncncaccgcc	420
agnacgcnc	tccgcccnc	gacgcccnn	cccgccgcgc	tcacctcat	ggncnncng	480
cccgcctcnc	ncnctgcnc	gccgncnngg	cgcgccgcc	cnnccgngtn	ccncncgngg	540
cccncgngn	angcngtgcg	cnnccngncc	gngccgncn	ncaccctccg	ncncccgcc	600
cgcgcgtgg	gggctccgc	cncgcgntc	antcccncc	cntncgccca	ctntccgntc	660
cnnctcnc	gctcngcng	cgcncncnc	ccccccc			697

<210> 18

<211> 670

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(670)

<223> n = A,T,C or G

<400> 18

ctcgtgtgaa	gggtgcagta	cctaagccgg	agcggggtag	aggcgggccg	gcacccccctt	60
ctgacctcca	gtgccgcgg	cctcaagatc	agacatggcc	cagaacttga	acgacttggc	120
gggacggctg	cccgccgggc	cccggggcat	gggcacggcc	ctgaagctgt	tgctgggggc	180
cggcgccgtg	gcctacgggtg	tgccggaatc	tgtgttcacc	gtggaaggcg	ggcncagagc	240
catcttcttc	aatcggtatc	gtggagtgc	caggacacta	tcctgggccc	anggccttca	300
cttcaggatc	cttggttcca	gtaccccanc	atctatgaca	ttcggggccag	acctcgaaaa	360
aatctcctcc	ctacagggtc	caaagaccta	cagatgggtga	atatctccct	gcgagtgttg	420
tctcgaccaa	tgctcangaa	cttctaaca	tggtccancg	cctaagggtc	ggactacnaa	480
gaacgantgt	tgccgtccat	tgctacgaag	tgctcaagaa	tttnggtggc	caagttcaat	540
gncctcann	ctgatnccc	agcggggcca	agttanccct	ggttgatccc	cgggganctg	600
acnnaaaagg	gccaaggact	ccccctcatc	ctggataatg	tggccntcac	aaagctcaac	660
tttanccacc						670

<210> 19

<211> 606

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(606)

<223> n = A,T,C or G

<400> 22

```

acaattttca ttatcttaag cacattgtac atttctacag aacctgtgat tattctcgca      60
tgataaggat ggtacttgca tatggtgaat tactactggt gacagtttcc gcagaaatcc      120
tatttcagtg gaccaacatt gtggcatggc agcaaagcc aacattttgt ggaatagcag      180
caaatctaca agagacctg gttggttttt cgttttgttt tctttgtttt tcccccttc      240
tctgaatca gcagggatgg aangagggtg gggaagttat gaattactcc ttccagtagt      300
agctctgaag tgtcacattt aatatcagtt ttttttaaac atgattctag ttnaatgtag      360
aagagagaag aaagaggaag tgttcacttt ttaatacac tgatttagaa atttgatgtc      420
ttatatcagt agttctgagg tattgatagc ttgctttatt tctgccttta cgttgacagt      480
gttgaagcag ggtgaataac taggggcata tatatttttt tttttgttaa gctgtttcat      540
gatgttttct ttggaatttc cggataagtt caggaaaaca tctgcatggt gttatctagt      600
ctgaagtten tatccatctc attacaacaa aaacncccag aacggnntg      649

```

```

<210> 23
<211> 669
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(669)
<223> n = A,T,C or G

```

```

<400> 23
actagtgccg tactggetga aatccctgca ggaccaggaa gagaaccagt tcagactttg      60
tactctcagt caccagctct ggaattagat aaattccttg aagatgtcag gaatgggac      120
tactctctga cagccttttg gctgcctcgg cccagcagc cacagcagga ggaggtgaca      180
tcacctgtcg tgcccccttc tgtcaagact ccgacacctg aaccagctga ggtggagact      240
cgcaagggtg tgctgatgca gtgcaacatt gagtcgggtg aggagggagt caaacaccac      300
ctgacacttc tgctgaagtt ggaggacaaa ctgaaccggc acctgagctg tgacctgatg      360
ccaaatgaga atatccccga gttggcggtt gagctggtgc agctgggctt cattagttag      420
gctgaccaga gccggttgac ttctctgcta gaagagactt gaacaagttc aattttgcc      480
ggaacagtac cctcaactca gccgtgtgca cgtctcctc ttagagctca ctcgggccag      540
gcctgatctt gcgctgtggc tgtcctggac gtgctgcacc ctctgtcctt cccccagtc      600
agtattacct gtgaagccct tccctccttt attattcagg anggctgggg gggctccttg      660
nttctaacc

```

```

<210> 24
<211> 442
<212> DNA
<213> Homo sapien

```

```

<400> 24
actagtacca tcttgacaga ggatacatgc tcccaaaacg tttgttacca cacttaaaaa      60
tcactgccat cattaagcat cagtttcaaa attatagcca ttcattgattt actttttcca      120
gatgactatc attattctag tcttttgaat ttgtaagggg aaaaaaaaca aaaacaaaaa      180
cttacgatgc actttttctcc agcacatcag atttcaaatt gaaaattaaa gacatgctat      240
ggtaatgcac ttgctagtac tacacacttt ggtacaacaa aaaacagagg caagaaacaa      300
cggaaagaga aaagccttcc tttgttggcc cttaaactga gtcaagatct gaaatgtaga      360
gatgatctct gacgatacct gtatgttctt atttgttaaa taaaattgct ggtatgaaat      420
gacctaaaaa aaaaaaaaga aa

```

```

<210> 25
<211> 656
<212> DNA

```

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(656)

<223> n = A,T,C or G

<400> 25

tgcaagtacc	acacactggt	tgaattttgc	acaaaaagtg	actgtaggat	caggtgatag	60
ccccggaatg	tacagtgtct	tggtgcacca	agatgccttc	taaaggctga	cataccttgg	120
accctaatagg	ggcagagagt	atagccctag	cccagtgggtg	acatgaccac	tcaccttggg	180
aggcctgagg	tagaggggag	tggtatgtgt	tttctcagtg	gaagcagcac	atgagtgggt	240
gacaggatgt	tagataaagg	ctctagttag	ggtgtcattg	tcatttgaga	gactgacaca	300
ctcctagcag	ctggtaaagg	ggtgctggan	gccatggagg	anctctagaa	acattagcat	360
gggctgatct	gattacttcc	tggtatcccg	ctcactttta	tggaagtct	tattagangg	420
atgggacagt	tttccatata	cttgctgtgg	agctctggaa	cactctctaa	atttccctct	480
attaaaaatc	actgccctaa	ctacacttcc	tccttgaagg	aatagaatg	gaactttctc	540
tgacatannt	cttggcatgg	ggagccagcc	acaaatgana	atctgaacgt	gtccagggtt	600
ctcctganac	tcctctacat	agaattgggt	aaacctccc	ttggaataag	gaaaaa	656

<210> 26

<211> 434

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(434)

<223> n = A,T,C or G

<400> 26

actagttcag	actgccacgc	caaccccaga	aaatacccca	catgccagaa	aagtgaagtc	60
ctaggtgttt	ccatctatgt	ttcaatctgt	ccatctacca	ggcctcgcca	taaaaacaaa	120
acaaaaaaaa	gctgccaggt	tttagaagca	gttctggtct	caaaaccatc	aggatcctgc	180
caccagggtt	cttttgaat	agtaccacat	gtaaaaggga	atttggcttt	cacttcatct	240
aataactgaa	ttgtcaggct	ttgattgata	attgtagaaa	taagtgcct	tctgttggtg	300
gaataagtta	taatcagtat	tcctctcttt	gttttttgtc	actcttttct	ctctaattgt	360
gtcatttgta	ctgtttgaaa	aatatttctt	ctatnaaatt	aaactaacct	gccttaaaaa	420
aaaaaaaaaa	aaaa					434

<210> 27

<211> 654

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(654)

<223> n = A,T,C or G

<400> 27

actagtccaa	cacagtcaga	aacattgttt	tgaatcctct	gtaaaccaag	gcattaatct	60
taataaacca	ggatccattt	aggtaccact	tgatataaaa	aggatatcca	taatgaatat	120
tttatactgc	atcctttaca	ttagccacta	aatacgttat	tgcttgatga	agacctttca	180

```
<210> 28
<211> 670
<212> DNA
<213> Homo sapien
```

<400>	28						
gcaca	tactgggagg	atttccacag	ctgcacggtc	acagccctta	cggattgcc		60
ggcg	aaagatatgt	gggataaact	gagaaaagaa	nccaaaaacc	tcaacatcca		120
gctta	ttcgaaactct	gcggcagcgg	caacggggcg	gcggggtccc	tgtctccggc		180
cgggt	ctcctggtgt	ctctctcggc	agcttttagc	acctgncctt	ccttctgagc		240
gccag	ctccccccgc	ggcgcccacc	cacnctcact	ccatgctccc	ggaaatcgag		300
gatca	ttagtctctt	ggggacgctt	gtgattctct	gtgatgctga	aaaacactca		360
ggaat	gtgggaaatc	ctganctctt	tnttatntcg	tntgatttct	tgtgttttat		420
aaaa	gttaccaatc	agtgaccaac	cnagcacagc	caaaaatcgg	acntcngctt		480
cgtct	tcacacacag	aataagaaaa	cggcaaacc	accccacttt	tnantttnat		540
ctaan	ttttttctgt	tgggcaaaaag	aatctcagga	acngccctgg	ggccnccgta		600
gtaaa	ccnagctagt	tncatgaaaa	atgatgggct	ccnctcaat	gggaaagcca		660
aaqnc							670

```
<210> 29
<211> 551
<212> DNA
<213> Homo sapien
```

```
<220>  
<221> misc_feature  
<222> (1)...(551)  
<223> n = A,T,C or G
```

<400> 29						
actagtcctc	cacagcctgt	gaatccccct	agacctttca	agcatagtg	gcgagagaaga	60
agatctcagc	gttttagccac	cttaccctatg	cctgatgatt	ctgtagaaaa	ggttttcttct	120
ccctctccag	ccactgatgg	gaaagtattc	tccatcagtt	ctcaaaatca	gcaagaatct	180
tcagtaccag	aggtgcctga	tgttgccacat	ttgccacttg	agaagctggg	accctgtctc	240
cctcttgact	taagtcgtgg	ttcagaagtt	acagcaccgg	tagcctcaga	ttcctctttac	300
cgtaatgaat	gtcccagggc	agaaaaagag	gatacncaga	tgcttccaaa	tccttctttcc	360
aaagcaatag	ctgatgggaa	gaggagctcc	agcagcagca	ggaatatcga	aaacagaaaa	420
aaaagtgaaa	ttgggaagac	aaaagctcaa	cagcatttgg	taaggagaaa	aganaagatg	480
aggaaggaag	agagaagaga	gacnaagatc	nctacggacc	gnnnccgaag	aagaagaagn	540
aaaaaaaaaa	a					551

```

<210> 30
<211> 684
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(684)
<223> n = A,T,C or G

```

```

<400> 30
actagttcta tctggaaaaa gcccggttg gaagaagctg tggagagtgc gtgtgcaatg      60
cgagactcat ttcttggaag catccctggc aaaaatgcag ctgagtacaa ggttatcact    120
gtgatagaac ctggactgct ttttgagata atagagatgc tgcagtctga agagacttcc    180
agcacctctc agttgaatga attaatgatg gcttctgagt caactttact ggctcaggaa    240
ccacgagaga tgactgcaga tgtaatcgag cttaaaggga aattcctcat caacttagaa    300
ggtggtgata ttcgtgaaga gtcttctctat aaagtaattg tcatgccgac tacgaaagaa    360
aaatgcccc  gttgttgga  gtatacagcg ggagtcttca gatacactgt gtctctgatg    420
tgcagaagtt gtcagtggga aaatagtatt aacagctcac tgcagcaaga accctcctga    480
cagtactggg ctagaagttt ggatggatta ttacaatat aggaaagaaa gccagaatt    540
aggtnatgag tggatgagta aatggtggan gatggggaat tcaaatacaga attatggaag    600
aagttnttcc tgttactata gaaaggaatt atgtttattt acatgcagaa aatatanatg    660
tgtggtgtgt accgtggatg gaan                                           684

```

```

<210> 31
<211> 654
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(654)
<223> n = A,T,C or G

```

```

<400> 31
gcgcagaaaa ggaaccaata tttcagaaac aagcttaata ggaacagctg cctgtacatc      60
aacatcttct cagaatgacc cagaagttat catcgtggga gctggcgtgc ttggctctgc    120
tttggcagct gtgctttcca gagatggaag aaaggtgaca gtcattgaga gagacttaaa    180
agagcctgac agaatagttg gagaattcct gcagccgggt ggttatcatg ttctcaaaga    240
ccttggtctt ggagatacag tggaaaggtc tgatgccag gttgtaaatg gttacatgat    300
tcatgatcag ggaaagcaaa tcagangttc agattcctta ccctctgtca gaaaacaatc    360
aagtgcagag tggaagagct ttccatcacg gaagattcat catgagtctc cggaaagcag    420
ctatggcaga gcccaatgca aagtttattg aaggtgttgt gttacagtta ttagagggaag    480
atgatgttgt gatgggagtt cagtacaagg ataaagagac tgggagatat caaggaaactc    540
catgctccac tgactgttgt tgcagatggg cttttctcca anttcaggaa aagcctggtc    600
tcaataaagt ttctgtatca ctcatcttggg tggcttctta tgaagaatgc nccc           654

```

```

<210> 32
<211> 673
<212> DNA
<213> Homo sapien

<220>

```

<400> 32

```
<210> 33
<211> 673
<212> DNA
<213> Homo sapien
```

```
<220>  
<221> misc_feature  
<222> (1)...(673)  
<223> n = A,T,C or G
```

<400> 33

```
<210> 34
<211> 684
<212> DNA
<213> Homo sapien
```

```
<220>
<221> misc_feature
<222> (1)...(684)
<223> n = A,T,C or G
```

<400> 34

actagtttat tcaagaaaag aacttactga ttctctgtt cctaaagcaa gagtggcagg 60


```

tgatcagggc tgggtgtagca tccggttcct ttagtgagc taactgcatt tgtcactgat 120
gaccaaggag gaaatcacta agacatttga gaagcagtgg tatgaacggt cttggacaag 180
ccacagttct gagccttaac cctgtagttt gcacacaaga acgagctcca cctccccttc 240
ttcaggagga atctgtgagg atagattggc tggacttttc aatgggttctg gggtgcaagt 300
gggcaactgtt atggctgggt atggagcggg cagccccagg aatcagagcc tcagcccggc 360
tgcttggttg gaaggtagag gtgttcagca ccttcggaaa aagggcataa agtngtgggg 420
gacaattctc agtccaagaa gaatgcattg accattgctg gctatttgct tncctagtan 480
gaattggatn catttttgac cangatnntt ctncatgct tntttgcaat gaaatcaaat 540
cccgcattat ctacaagtgg tatgaagtc tgcnncccc agagaggctg ttcaggcnat 600
gtcttccaag ggcagggtgg gttacacat tttacctccc ctctcccccc agattatgna 660
cncagaagga atttntttcc tccc 684

```

```

<210> 35
<211> 614
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(614)
<223> n = A,T,C or G

```

```

<400> 35
actagtccaa cgcgttngcn aatattcccc tggtagccta ctcccttacc cccgaatatt 60
ggtaagatcg agcaatggct tcaggacatg ggttctcttc tctgtgatc attcaagtgc 120
tcaactgatg aagactggct tgtctcagtg tntcaacctc accagggtg tctcttggtc 180
cacacctcgc tccctgttag tgcgtagtga cagcccccat canatgacct tggccaagtc 240
acggtttctc tgtgggtcaat gttggtnggc tgattgggtg aaagtanggt ggaccaaagg 300
aagncncgtg agcagncanc nccagttctg caccagcagc gcctccgtcc tactnggggtg 360
ttcengtttc tcttgggcct gngtgggcta nggectgatt cgggaanatg cctttgcang 420
gaaggganga taantgggat ctaccaattg attctggcaa aacnatntct aagattnttn 480
tgctttatgt ggganacana tctantcttc atttntgtct gnatanaca ccctactcgt 540
gntcgancnc gtcttcgatt ttcgganaca cnccantnaa tactggcggt ctgttggtta 600
aaaaaaaaaa aaaa 614

```

```

<210> 36
<211> 686
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(686)
<223> n = A,T,C or G

```

```

<400> 36
gtggctggcc cggttctccg cttctcccca tcccctactt tctccctcc ctccttttcc 60
ctccctcgtc gactgttgct tgetggtoe gaactccctg accctccct caccctccc 120
taacctcggg gccaccgat tgcccttctt ttctgttgcc ccagcccagc cctagtgtca 180
gggcgggggc ctggagcagc ccgaggcact gcagcagaag anaaaaaga cagcagcaac 240
ctcagctcgc cagtccggtc gctngcttcc cgccgcatgg caatnagaca gacgccgctc 300
acctgctctg ggcacacgag acccgtgggt gatttggcct tcagtggcat cacccttatg 360
ggtatttctt aatcagcgct tgcaaagatg gttaacctat gctacgccag ggagatacag 420
gagactggat tggaaacattt ttgggggtcta aaggtctgtt tggggtgcaa cactgaataa 480

```

```

ggatgccacc aaagcagcta cagcagctgc agatttcaca gccaagtgt gggatgctgt 540
ctcagganat naattgataa cctggctcat aacacattgt caagaatgtg gatttcccca 600
ggatattatt atttgtttac cggggganag gataactgtt tcnctatatt taattgaaca 660
aactnaaaca aaanctaagg aaatcc 686

```

```

<210> 37
<211> 681
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(681)
<223> n = A,T,C or G

```

```

<400> 37
gagacanacn naacgtcang agaanaaaag angcatggaa cacaanccag gcncgatggc 60
caccttccca ccagcancca gcgcccccca gcngccccc ngnccggang accangactc 120
cancctgnat caatctganc tctattcctg gccatnccet acctcggagg tggangccgn 180
aaaggtcgca cnnncagaga agctgctgcc ancaccancc gcccnnccc tgnccgggctn 240
nataggaaac tggtgaccnn gctgcanaat tcatacagga gcacgcgang ggcacnnnct 300
cacactgagt tnnngatgan gctnaccan ggacctnccc cagcnnattg annacnggac 360
tgccggaggaa ggaagacccc gnaacggatc ctggccggcn tgccaccccc ccacccctag 420
gattatnccc cttgactgag tctctgagg gctacccgaa cccgctcca ttccctacca 480
natnntgctc natcgggact gacangctgg ggatnggagg ggctatcccc cancatcccc 540
tnanaccaac agcnaengan natnggggct cccnggggtc ggngcaacnc tctncaccc 600
cggcgcnggc cttcggtgnt gtctctcctc aacnaattcc naaanggcgg gccccccngt 660
ggactcctcn ttgttccctc c 681

```

```

<210> 38
<211> 687
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(687)
<223> n = A,T,C or G

```

```

<400> 38
canaaaaaaa aaaacatggc cgaaaccagn aagctgcgcg atggcgccac ggcccctctt 60
ctcccgccct gtgtccggaa ggtttccctc cgaggcgccc cggtcccgcc aagcggagga 120
gaggcgggga cntgcccggg ccggagctca naggccctgg ggccgctctg ctctcccgcc 180
atcgcaaggg cggcgctaac ctnaggcctc cccgcaaagg tcccnhangc gngggcggcg 240
gggggctgtg anaaccgcaa aaanaacgct gggcgcgcn ggaacccgtc ccccccgcg 300
aaggananac ttccacagan gcagcgtttc cacagccan agccacnttt ctaggggtgat 360
gcacccagat aagttcctgn cggggaagct caccgctgtc aaaaaanctc ttcgctccac 420
cggcgcacna aggggangan ggcangangc tgcgcgccgc acaggtcatc tgatcacgtc 480
gcccgccta ntctgtttt gtgaatctcc actttgttca accccacccg ccgttctctc 540
ctccttgcc cttcctctna ccttaanaac cagcttctc taccnatng tanttctct 600
gcncnngtng aaattaattc ggctcncggg aacctcttnc ctgtggcaac tgctnaaaga 660
aactgctgtt ctgnntactg cngtccc 687

```

```

<210> 39

```

<211> 695
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(695)
 <223> n = A,T,C or G

<400> 39

actagtctgg	cctacaatag	tgtgattcat	gtaggacttc	tttcatcaat	tcaaaacccc	60
tagaaaaacg	tatacagatt	atataagtag	ggataagatt	tctaacattt	ctgggctctc	120
tgaccctctg	gctagactgt	ggaaaggagg	tattattata	gtatacaaca	ctgctgttgc	180
cttattagtt	ataacatgat	aggtgctgaa	ttgtgattca	caatttaaaa	acactgtaat	240
ccaaactttt	ttttttaact	gtagatcatg	catgtgaatg	ttaatgttaa	tttgttcaan	300
gttgttatgg	gtagaaaaaa	ccacatgcct	taaaatttta	aaaagcaggg	cccaaactta	360
ttagtttaaa	attaggggta	tgtttccagt	ttgttattaa	ntggttatag	ctctgtttag	420
aanaaatcna	ngaacangat	ttngaaantt	aagntgacat	tatttnccag	tgacttgtaa	480
atttgaaatc	anacacggca	ccttcctgtt	tggttctatt	ggnttttgaa	tccaancngg	540
ntccaaatct	tnttggaaac	ngtccnttta	acttttttac	nanatcttat	ttttttattt	600
tggaatggcc	ctattttaang	ttaaaagggg	ggggnnccac	naccattcnt	gaataaaaact	660
naatatatat	ccttgggtccc	ccaaaattta	aggng			695

<210> 40
 <211> 674
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(674)
 <223> n = A,T,C or G

<400> 40

actagttagtc	agttgggagt	ggttgctata	ccttgacttc	atttatatga	atttccactt	60
tattaaataa	tagaaaagaa	aatcccgggtg	cttgacgtag	agttatagga	cattctatgc	120
ttacagaaaa	tatagccatg	attgaaatca	aatagtaaag	gctgttctgg	ctttttatct	180
tcttagctca	tcttaaataa	gtagtacact	tgggatgcag	tgcgctctgaa	gtgctaataa	240
gttgtaacaa	tagcacaaat	cgaacttagg	atgtgtttct	tctcttctgt	gtttcgattt	300
tgatcaattc	tttaattttg	ggaacctata	atacagtttt	cctattcttg	gagataaaaa	360
ttaaatggat	cactgatatt	taagtcattc	tgcttctcat	ctnaatatcc	catattctgt	420
attagganaa	antacctccc	agcacagccc	cctctcaaac	cccacccaaa	accaagcatt	480
tggaatgagt	ctcctttatt	tccgaantgt	ggatgggtata	acccatatcn	ctccaatttc	540
tgnttgggtt	gggtattaat	ttgaactgtg	catgaaaagn	ggnaatcttt	nctttgggtc	600
aaanttttnc	ggttaatttg	nctngncaaa	tccaatttnc	tttaagggtg	tctttataaa	660
atttgctatt	cngg					674

<210> 41
 <211> 657
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature

<222> (1)...(657)

<223> n = A,T,C or G

<400> 41

gaaacatgca	agtaccacac	actgtttgaa	ttttgcacaa	aaagtgactg	tagggatcag	60
gtgatagccc	cggaatgtac	agtgtcttgg	tgcaccaaga	tgccttctaa	aggctgacat	120
accttggggac	cctaattggg	cagagagtat	agccctagcc	cagtgggtgac	atgaccactc	180
cctttggggag	gctgaagtta	aagggaatgg	tatgtgtttt	ctcatggaag	cagcacatga	240
atnggtnaca	ngatgttaaa	ntaaggntct	antttgggtg	tcttgtcatt	tgaaaaantg	300
acacactcct	ancanctggg	aaaggggtgc	tggaagccat	ggaagaactc	taaaaacatt	360
agcatgggct	gatctgatta	cttcctggca	tcccgtcac	ttttatggga	agtcttatta	420
naaggatggg	ananttttcc	atatccttgc	tgttggaact	ctggaacact	ctctaaattt	480
ccctctatta	aaaatcactg	nccttactac	acttcctcct	tganggaata	gaaatggacc	540
tttctctgac	ttagttcttg	gcatgggganc	cagcccaaat	taaaatctga	cttntccggt	600
ttctccngaa	ctcacctact	tgaattggta	aaacctcctt	tggaattagn	aaaaacc	657

<210> 42

<211> 389

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(389)

<223> n = A,T,C or G

<400> 42

actagtgtctg	aggaatgtaa	acaagtttgc	tgggccttgc	gagacttcac	caggttgttt	60
cgatagctca	cactcctgca	ctgtgcctgt	caccagga	tgtctttttt	aattagaaga	120
caggaagaaa	acaaaaacca	gactgtgtcc	cacaatcaga	aacctccgtt	gtggcagang	180
ggccttcacc	gccaccagg	tgtcccgcca	gacagggaga	gactccagcc	ttctgaggcc	240
atcctgaaga	attcctgttt	gggggtgtg	aaggaaaatc	acccggttt	aaaaagatgc	300
tgttgccctgc	ccgcgtngtn	gggaaggac	tggtttccctg	gtgaatttct	taaaagaaaa	360
atattttaag	ttaagaaaaa	aaaaaaaaa				389

<210> 43

<211> 279

<212> DNA

<213> Homo sapien

<400> 43

actagtgaca	agctcctggg	cttgagatgt	cttctcggtta	aggagatggg	ccttttggag	60
gtaaaggata	aaatgaatga	gttctgtcat	gattcactat	tctagaactt	gcatgacctt	120
tactgtgtta	gctctttgaa	tggtcttgaa	attttagact	ttctttgtaa	acaaataata	180
tgctccttatc	attgtataaa	agctgttatg	tgcaacagtg	tggagatcct	tgtctgattt	240
aataaaatac	ttaaactg	aaaaaaaaa	aaaaaaaaa			279

<210> 44

<211> 449

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(449)
 <223> n = A,T,C or G

<400> 44

actagtagca	tcttttctac	aacgttaaaa	ttgcagaagt	agcttatcat	taaaaaaciaa	60
caacaacaac	aataacaata	aatcctaagt	gtaaatcagt	tattctaccc	cctaccaagg	120
atatcagcct	gttttttccc	ttttttctcc	tgggaataat	tgtgggcttc	ttcccaaatt	180
tctacagcct	ctttctctct	ctcatgcttg	agcttccttg	tttgacgca	tgcgttgctg	240
aagantgggc	tgtttngctt	ggantncggt	ccnagtggaa	ncatgctttc	ccttgttact	300
gttggaagaa	actcaaacct	tcnancctta	ggtgttncca	ttttgtcaag	tcatacactgt	360
atttttgtac	tggcattaac	aaaaaaagaa	atnaaatatt	gttccattaa	actttaataa	420
aactttaaaa	gggaaaaaaa	aaaaaaaa				449

<210> 45
 <211> 559
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(559)
 <223> n = A,T,C or G

<400> 45

actagtgtgg	gggaatcacg	gacacttaaa	gtcaatctgc	gaaataattc	ttttattaca	60
cactcactga	agtttttgag	tcccagagag	ccattctatg	tcaaacattc	caagtactct	120
ttgagagccc	agcattacat	caacatgccc	gtgcagttca	aaccgaagtc	cgcaggcaaa	180
tttgaagctt	tgcttgctat	tcaaacagat	gaaggcaaga	gtattgctat	tcgactaatt	240
ggtgaagctc	ttggaaaaaa	ttnactagaa	tactttttgt	gttaagttaa	ttacataagt	300
tgtattttgt	taactttatc	tttctacact	acaattatgc	ttttgtatat	atattttgta	360
tgatggatat	ctataattgt	agattttggt	tttacaagct	aatactgaag	actcgactga	420
aatattatgt	atctagccca	tagtattgta	cttaactttt	acagggtgaa	aaaaaaattc	480
tgtgtttgca	ttgattatga	tattctgaat	aaatatggga	atatatttta	atgtgggtaa	540
aaaaaaaaaa	aaaaaggaa					559

<210> 46
 <211> 731
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(731)
 <223> n = A,T,C or G

<400> 46

actagttcta	gtaccatggc	tgtcatagat	gcaaccatta	tattccattt	agttttcttc	60
tcaggttccc	taacaattgt	ttgaaactga	atatatatgt	ttatgtatgt	gtgtgtgttc	120
actgtcatgt	atatggtgta	tatgggatgt	gtgcagtttt	cagttatata	tatattcata	180
tatacatatg	catatatatg	tataatatac	atatatacat	gcatacactt	gtataatata	240
catatatata	cacatatatg	cacacataatn	atcactgagt	tccaaagtga	gtctttatit	300
ggggcaattg	tattctctcc	ctctgtctgc	tcactgggct	tttgcaagac	atagcaattg	360
cttgatttcc	tttgataag	agtccttatct	tcggcactct	tgactctagc	cttaacttta	420
gatttctatt	ccagaatacc	tctcatatct	atcttaaaac	ctaaganggg	taaagangtc	480

ataagattgt	agtatgaaag	antttgctta	gttaaattat	atctcaggaa	actcattcat	540
ctacaaatta	aattgtaaaa	tgatggtttg	ttgtatctga	aaaaatgttt	agaacaagaa	600
atgtaactgg	gtacctgtta	tatcaaagaa	cctcnattta	ttaagtctcc	tcatagccan	660
atccttatat	ngccctctct	gacctgannt	aatananact	tgaataatga	atagttaatt	720
taggnntggg	c					731

<210> 47
 <211> 640
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(640)
 <223> n = A,T,C or G

<400> 47

tgcgngccgg	tttggccctt	ctttgtanga	cactttcatc	cgcctgaaa	tcttcccgat	60
cgtaataaac	tcctcaggtc	cctgcctgca	cagggttttt	tcttantttg	ttgcctaaca	120
gtacaccaa	tgtgacatcc	tttcaccaat	atngatttct	tcataccaca	tcntcnatgg	180
anacgactnc	aacaattttt	tgatnaccn	aaanactggg	ggctnnaana	agtacantct	240
ggagcagcat	ggacctgtcn	gcnactaang	gaacaanagt	nntgaacatt	tacacaacct	300
ttgggtatgtc	ttactgaaag	anagaaacat	gcttctnncc	ctagaccacg	aggncaacccg	360
caganattgc	caatgccaa	tcgagcggt	tagatcagg	aatacattcc	atggatgcat	420
tacatacnnt	gtccccgaaa	nanaagatgc	cctaanggct	tcttcanact	ggccngaaa	480
acantacac	ctggtgcttg	ganaacanan	tctttggaag	atcatctggc	acaagttccc	540
cccagtggtg	tttnccttgg	cacctanctt	accanatcna	ttcggaancc	attctttgcc	600
ntggcnttnt	ntgggacca	ntcttctcac	aactgnaccc			640

<210> 48
 <211> 257
 <212> DNA
 <213> Homo sapien

<400> 48

actagtatat	gaaaatgtaa	atatcacttg	tgtactcaaa	caaaagttgg	tcttaagctt	60
ccaccttgag	cagccttgga	aacctaacct	gcctctttta	gcataatcac	attttctaaa	120
tgattttctt	tgttctcgaa	aaagtgtatt	gtattagttt	tacatttggt	ttttggaaga	180
ttatatttgt	atatgtatca	tcataaaata	tttaaataaa	aagtatcttt	agagtgaaaa	240
aaaaaaaaaa	aaaaaaaa					257

<210> 49
 <211> 652
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(652)
 <223> n = A,T,C or G

<400> 49

actagttcag	atgagtggct	gtggaagggg	cccccttgtc	attttcatta	taaccaatt	60
tccacttatt	tgaactctta	agtcataaat	gtataatgac	ttatgaatta	gcacagttaa	120

```
<210> 50
<211> 650
<212> DNA
<213> Homo sapien
```

<400> 50						
ctttg	attttttttag	ggcttgtgcc	ctgttttctact	tataggggtct	agaatgcttg	60
agtaa	aaaggagatg	cccaatatct	aaagctgcta	aatgttctct	ttgccataaa	120
cgtgt	aactgtgtga	acacttgagg	tttttctcct	ctgtcccgag	gtcgtcgtct	180
ctttt	ttgggttctt	tctagaagat	tgagaaatgc	atatgacagg	ctgagancac	240
caaac	acacaagctc	tcagccacan	gcagcttctc	cacagcccca	gcttcgcaca	300
ctgga	nggctgcctg	ggggaggcag	acatgggagt	gccaaggtgg	ccagatggtt	360
actac	aatgtcttta	tttttaactg	tttgccactg	ctgccctcac	ccctgcccg	420
gagta	ccgtctgcc	canacaagtg	ggantgaaat	gggggtgggg	gggaacactg	480
cantt	aggggggtgc	taactgaaca	gtagggatan	aaggtgtgaa	cctgngaant	540
tataa	attatnttcc	ttgttanatt	tatttttttaa	tttaactctc	gttnaactgc	600
gaaaa	ggggaaaaaaa	aaaaaaaaaat	tctnttttaa	cacatgaaca		650

```
<210> 51
<211> 545
<212> DNA
<213> Homo sapien
```

```
<220>
<221> misc_feature
<222> (1)...(545)
<223> n = A,T,C or G
```

<400>	51						
tggcgtgcaa	ccagggtage	tgaagtttgg	gtctgggact	ggagattggc	cattaggcct		60
cctganattc	cagctccctt	ccaccaagcc	cagtcttgct	acgtggcaca	gggcaaacct		120
gactcccttt	gggcctcagt	ttccctctcc	cttcatgana	tgaaaagaat	actacttttt		180
cttgttggtc	taacnttget	ggacncaaag	tgtngtcatt	attgttgtat	tgggtgatgt		240
gtncaaaact	gcagaagctc	actgcctatg	agaggaanta	agagagatag	tggatganag		300
ggacanaagg	agtcattatt	tggatatagat	ccaccntcc	caacctttct	ctcctcagtc		360
cctgcnccctc	atgtntctgg	tntgggtgagt	ccttttgtgcc	accanccatc	atgcttttgc		420
ttgctgccat	cctgggaagg	gggtgnatcg	tctcacaact	tgtttgtcatc	gtttganatg		480
catgctttct	tnatnaaaca	aanaaannaa	tgtttgacag	ngtttaaaat	aaaaaanaaa		540
caaaa							545

<210> 52
 <211> 678
 <212> DNA
 <213> Homo sapien

 <220>
 <221> misc_feature
 <222> (1)...(678)
 <223> n = A,T,C or G

<400> 52
 actagtagaa gaactttgcc gcttttgtgc ctctcacagg cgcctaaagt cattgccatg 60
 ggaggaagac gatttggggg gggagggggg gggggcangg tccgtggggc ttccctant 120
 ntatctccat ntccantgnn cnntgtcgcc tcttccctcg tncattnga anttantccc 180
 tggnccccnn nccctctccn nctnccnct ccccccctcg nccctccnn cttttntan 240
 nettcccat ctccntcccc cctnanngtc ccaacnccgn cagcaatnnc ncacttnctc 300
 netccnccc tccnnccgtt cttctnttct cnaentntnc ncnntnccn tgcnnntnaa 360
 annctctccc cnetgcaanc gattctctcc ctcennnan ctntccactc cntncttctc 420
 ncnctctcc nttctcnnc ccacctcten ccttcgnccc cantacnctc nccnccctn 480
 cgnntcnttn nnntctcnn accncccncc tcccttncce cctcttctcc ccggtntntc 540
 tetctccnc ncnccnccct cncnccntcc nngcgncnt ttcgcgccn cncnccntt 600
 ccttctcnc cantccatcn cntntnccat netnccncc nctcacnccc gctnccccc 660
 ntctctttca cacngtcc 678

<210> 53
 <211> 502
 <212> DNA
 <213> Homo sapien

 <220>
 <221> misc_feature
 <222> (1)...(502)
 <223> n = A,T,C or G

<400> 53
 tgaagatcct ggtgtcgcca tgggcgcgcg cccgcgccgt tgttaccggt attgtaagaa 60
 caagccgtac ccaaagtctc gcttctgcgc aggtgtccct gatgccaaaa ttgcgatttt 120
 tgacctgggg cggaaaaang caaaantgga tgagtctccg ctttgtggcc acatggtgtc 180
 agatcaatat gagcagctgt cctctgaagc cctgnanct gcccgaattt gtgccaataa 240
 gtacatggta aaaagtngtg gnaagatgc ttccatatcc ggggtgcggnt ccaccccttc 300
 cacgtcatcc gcatcaacaa gatgttgctc tgtgctgggg ctgacagget cccaacaggc 360
 atgcgaagtg cctttggaaa acccanggca ctgtggccag gggttcacatt gggccaattn 420
 atcatgttca tccgcaccaa ctgcagaaca angaantgt naattnaagc cctgcccagg 480
 gncaanttca aatttcccgg cc 502

<210> 54
 <211> 494
 <212> DNA
 <213> Homo sapien

 <220>
 <221> misc_feature
 <222> (1)...(494)

<223> n = A,T,C or G

<400> 54

```
actagtccaa gaaaaaatatg cttaaatgtat attacaaagg ctttgtatat gttaacctgt      60
tttaatgccaa aaagttttgct ttgtccacaaa ttctcttaag acctcttcag aaagggattt     120
gtttgcctta atgaatactg ttgggaaaaaa acacagtata atgagtgaag agggcagaag     180
caagaaatct ctacatctta ggcactccaa gaagaatgag tatccacatt tagatggcac     240
attatgagga ctttaatctt tccttaaaca caataatgtt ttcttttttc ttttattcac     300
atgatttcta agtatatttt tcatgcagga cagtttttca accttgatgt acagtgactg     360
tgttaaatct ttctttcagt ggcaacctct ataatcttta aaatatgggt agcatcttgt     420
ctgttttgaa ngggatatga cnatnaatct atcagatggg aaatcctgtt tccaagttag     480
aaaaaaaaaa aaaa
```

<210> 55

<211> 606

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(606)

<223> n = A,T,C or G

<400> 55

```
actagtaaaa agcagcattg ccaaataatc cctaattttc cactaaaaat ataatgaaat      60
gatgttaagc tttttgaaaa gtttaggtta aacctactgt tgttagatta atgtatttgt     120
tgcttccctt tatctggaat gtggcattag cttttttatt ttaaccctct ttaattctta     180
ttcaattcca tgacttaagg ttggagagct aaacactggg atttttggat aacagactga     240
cagttttgca taattataat cggcattgta catagaaagg atatggctac cttttgttaa     300
atctgcactt tctaaatatt aaaaaagggg aatgaagtat aaatcaattt ttgtataatc     360
tgtttgaaac atgantttta tttgcttaat attanggett tgcccttttc tgttagtctc     420
ttgggatect gtgtaaaact gttctcatta aacaccaaac agttaagtec attctctggt     480
actagctaca aattccggtt catattctac ntaacaattt aaattaactg aaatatttct     540
anatggtcta cttctgtcnt ataaaaacna aacttgantt nccaaaaaaa aaaaaaaaaa     600
aaaaaa
```

<210> 56

<211> 183

<212> DNA

<213> Homo sapien

<400> 56

```
actagtatat ttaaacttac aggcttattt gtaatgtaaa ccaccatttt aatgtactgt      60
aattaacatg gttataatac gtacaatcct tccctcatcc catcacacaa ctttttttgt     120
gtgtgataaa ctgattttgg tttgcaataa aaccttgaaa aataaaaaaa aaaaaaaaaa     180
aaa
```

<210> 57

<211> 622

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(622)

<223> n = A,T,C or G

<400> 57

```
actagtcact actgtcttct ccttgtagct aatcaatcaa tattcttccc ttgcctgtgg      60
gcagtgaggaga gtgctgctgg gtgtacgctg cacctgcccc ctgagttggg gaaagaggat      120
aatcagtgag cactgttctg ctcagagctc ctgatctacc ccacccccta ggatccagga      180
ctgggtcaaaa gctgcatgaa accaggccct ggcagcaacc tgggaatggc tggaggtggg      240
agagaacctg acttctcttt cctctccct cctccaacat tactggaact ctatcctgtt      300
agggatcttc tgagcttggt tccctgctgg gtgggacaga agacaaagga gaagggangg      360
tctacaanaa gcagcccttc tttgtcctct ggggttaatg agcttgacct ananttcag      420
gaganaccan aagcctctga tttttaattt cctnnaaatg tttgaagtnt atatntacat      480
atatatatatt ctttnaatnt ttgagtcttt gatatgtctt aaaatccant ccctctgccn      540
gaaacctgaa ttaaaacat gaanaaaaat gtttncctta aagatgttan taattaattg      600
aaacttgaaa aaaaaaaaaa aa                                     622
```

<210> 58

<211> 433

<212> DNA

<213> Homo sapien

<400> 58

```
gaacaaaattc tgattgggta tgtaccgtca aaagacttga agaaatttca tgattttgca      60
gtgtggaagc gttgaaaatt gaaagttact gcttttccac ttgtcatat agtaaaggga      120
tcctttcagc tgccagtgtt gaataatgta tcatccagag tgatgttatc tgtgacagtc      180
accagcttta agctgaacca ttttatgaat accaaataaa tagacctctt gtactgaaaa      240
catatttggt actttaatcg tgctgcttgg atagaaatat ttttactggg tcttctgaat      300
tgacagtaaa cctgtccatt atgaatggcc tactgttcta ttatttggtt tgacttgaat      360
ttatccacca aagacttcat ttgtgtatca tcaataaagt tgtatgtttc aactgaaaaa      420
aaaaaaaaaaa aaa                                           433
```

<210> 59

<211> 649

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(649)

<223> n = A,T,C or G

<400> 59

```
actagttatt atctgacttt cnggttataa tcattctaag gagtgtgaag tagcctctgg      60
tgtcatttgg atttgcatth ctctgatgag tgatgctatc aagcaccttt gctgggtgctg      120
ttggccatat gtgtatgttc cctggagaag tgtctgtgct gagccttggc ccacttttta      180
attaggcgtn tgtcttttta ttactgagtt gtaaganttc tttatatatt ctggattcta      240
gacccttatc agatacatgg tttgcaaata ttttctccca ttctgtgggt tgtgttttca      300
ctttatcgat aatgtcctta gacatataat aaatttgtat tttaaaagtg acttgatttg      360
ggctgtgcaa ggtgggctca cgcttgtaat ccagcactt tgggagactg aggtgggtgg      420
atcatatgan gangctagga gtctgaggtc agcctggcca gcatagcgaa aacttgtctc      480
tacnaaaaat acaaaaatta gtcaggcatg gtggtgcacg tctgtaatac cagcttctca      540
ggangctgan gcacaaggat cacttgaacc ccagaangaa gangttgcag tganctgaag      600
atcatgccag ggcaacaaaa atgagaactt gtttaaaaaa aaaaaaaaaa      649
```

<210> 60
 <211> 423
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(423)
 <223> n = A,T,C or G

<400> 60

actagttcag	gccttccagt	tcaactgacaa	acatggggaa	gtgtgcccag	ctggctggaa	60
acctggcagt	gataccatca	agcctgatgt	ccaaaagagc	aaagaatatt	tctccaagca	120
gaagtgagcg	ctgggctgtt	ttagtgccag	gctgcggttg	gcagccatga	gaacaaaacc	180
tcttctgtat	tttttttttc	cattagtana	acacaagact	cngattcagc	cgaattgtgg	240
tgtcttacaa	ggcagggttt	tcctacaggg	ggtgganaaa	acagcctttc	ttcctttggt	300
aggaatggcc	tgagttggcg	ttgtgggcag	gctactgggt	tgtatgatgt	attagtagag	360
caaccatta	atcttttgta	gtttgtatna	aacttganct	gagaccttaa	acaaaaaaaa	420
aaa						423

<210> 61
 <211> 423
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(423)
 <223> n = A,T,C or G

<400> 61

cgggactgga	atgtaaagtg	aagttcggag	ctctgagcac	gggtctcttc	cgccgggtcc	60
tccttcccc	gaccccagag	ggagaggccc	accccgcccc	gccccgcccc	agcccctgct	120
caggtctgag	tatggctggg	agtcgggggc	cacaggcctc	tagctgtgct	gctcaagaag	180
actggatcag	ggtanctaca	agtggccggg	ccttgccctt	gggattctac	cctgttccta	240
atttgggtgt	ggggtgcggg	gtccctggcc	cccttttcca	cactncctcc	ctcngacag	300
caacctccct	tggggcaatt	gggcctggnt	ctccncccg	tgttgcnaac	ctttgttggt	360
ttaaggncct	taaaaatggt	annttttccc	ntgcnggggt	taaaaaagga	aaaaactnaa	420
aaa						423

<210> 62
 <211> 683
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(683)
 <223> n = A,T,C or G

<400> 62

gctggagagg	ggtacggact	ttcttggagt	tgtcccaggt	tggaatgaga	ctgaactcaa	60
gaagagacc	taagagactg	gggaatgggt	cctgccttca	ggaaagtga	agacgcttag	120
gctgtcaaca	cttaaaggaa	gtccccttga	agcccagagt	ggacagacta	gacccattga	180

tggggccact	ggccatggtc	ogtggacaag	acattccngt	gggcatggc	acaccggggg	240
ggatcaaaat	gtgtacttgt	ggggtctcgc	cccttgccaa	aaccaaacca	ntcccactcc	300
tgtcnttgga	ctttcttccc	attccctcct	ccccaaatgc	acttcccctc	ctcccctctgc	360
ccctcctgtg	tttttggaat	tctgtttccc	tcaaaattgt	taatttttta	nttttngacc	420
atgaacttat	gtttggggtc	nangttcccc	ttaccaatgc	atactaatat	attaatgggt	480
atttatTTTT	gaaatatTTT	ttaatgaact	tggaaaaaat	tnntggaatt	tccttncttc	540
cnTTTTnttt	gggggggggtg	gggggntggg	ttaaaatTTT	tttggaancc	cnatnggaaa	600
ttnttacttg	gggccccct	naaaaaantn	antccaatt	cttnnatngc	ccctnttccn	660
ctaaaaaaaa	ananannaaa	aan				683

<210> 63
 <211> 731
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(731)
 <223> n = A,T,C or G

<400> 63						
actagtcata	aagggtgtgc	gogtcttcga	cgtggcggtc	ttggcgccac	tgctgcgaga	60
cccggccctg	gacctcaagg	tcateccactt	ggtgcgtgat	ccccgcgcgg	tggcgagttc	120
acggatccgc	tcgcgccacg	gcctcatccg	tgagagccta	caggtggtgc	gcagccgaga	180
ccgcgagctc	acgcgatgcc	cttcttgtag	gccgcggggc	acaagcttgg	cgcccanaaa	240
gaaggcgtng	ggggcccgca	aantaccacg	ctctgggcgc	tatggaangt	cctcttgcaa	300
taatattggt	tnaaaanctg	canaanagcc	cctgcancct	cctgaactgg	gntgcagggc	360
cncttacctn	gtttggntgc	ggttacaaag	aacctgtttt	ggaaaaccct	nccnaaaacc	420
ttccgggaaa	attntncaaa	ttttntttgg	ggaattnttg	ggtaaaccct	ccnaaaatgg	480
gaaacntttt	tgccctnnaa	antaaaccat	tnggttcggg	gggccccccc	ncaaaaccct	540
ttttnttttt	ttntgcccc	cantnncccc	ccggggcccc	tttttttngg	ggaaaaancc	600
ccccctncc	nanantttta	aaaggngggg	anaatttttt	nttncccccc	gggncccccn	660
gngntaaaa	nggtttcncc	cccccgaggg	gnggggnnnc	ctcnnaaacc	cntntcnna	720
ccncttttt	n					731

<210> 64
 <211> 313
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(313)
 <223> n = A,T,C or G

<400> 64						
actagttgtg	caaaccacga	ctgaagaaag	acgaaaagtg	ggaaataact	tgcaacgtct	60
gtagagatg	gttgctacac	atgttgggtc	tgtagagaaa	catcttgagg	agcagattgc	120
taaagttgat	agagaatatg	agaatgcat	gtcagaagat	ctctcggaag	atattaaaga	180
gattagagat	aagtatgaga	agaaagctac	tctaattaag	tcttctgaag	aatgaagatn	240
aaatgttgat	catgtatata	tatccatagt	gaataaaatt	gtctcagtaa	agttgtaaaa	300
aaaaaaaaaa	aaa					313

<210> 65

<211> 420
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(420)
 <223> n = A,T,C or G

<400> 65

actagttccc	tggcaggcaa	gggcttccaa	ctgaggcagt	gcatgtgtgg	cagagagagg	60
caggaagctg	gcagtggcag	cttctgtgtc	tagggagggg	tgtggctccc	tccttccctg	120
tctgggaggt	tggaggggaag	aatctaggcc	ttagcttgcc	ctcctgccac	ccttcccctt	180
gtagatactg	ccttaacact	ccctcctctc	tcagctgtgg	ctgccaccca	agccaggttt	240
ctccgtgtc	actaatat	ttccaggaaa	ggtgtgtgga	agacatgagc	cgtgtataat	300
atttgtttta	acattttcat	tgcaagtatt	gaccatcacc	cttggttgtg	tatcgttgta	360
acacaaatta	atgatattaa	aaagcatcca	aacaaagccn	annnnnaana	nnannngaaa	420

<210> 66
 <211> 676
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(676)
 <223> n = A,T,C or G

<400> 66

actagtttcc	tatgatcatt	aaactcattc	tcagggttaa	gaaaggaatg	taaattttctg	60
cctcaatttg	tacttcatca	ataagttttt	gaagagtgc	gatttttagt	caggtcttaa	120
aaataaaact	acaaatctgg	atgcatttct	aaattctgca	aatgtttcct	ggggtgactt	180
aacaaggaat	aatcccacaa	tatacctagc	tacctaatac	atggagctgg	ggctcaaccc	240
actgttttta	aggatttgcg	cttacttgtg	gctgaggaaa	aataagtagt	tccgagggaa	300
gtagttttta	aatgtgagct	tatagatngg	aaacagaata	tcaacttaat	tatggaaatt	360
gttagaaacc	tgttctcttg	ttatctgaat	cttgattgca	attactattg	tactggatag	420
actccagccc	attgcaaagt	ctcagatata	ttanctgtgt	agttgaattc	cttggaaatt	480
ctttttaaga	aaaaattgga	gtttnaaaga	aataaacccc	tttgttaaat	gaagcttggc	540
tttttggtga	aaaanaatca	tcccgtaggg	ottattgttt	aaaaanggaa	ttttaagcct	600
ccctggaaaa	anttgttaat	taaaggggga	aaatgntggg	naaaaattat	ccgttagggg	660
ttaaagggaa	aactta					676

<210> 67
 <211> 620
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(620)
 <223> n = A,T,C or G

<400> 67

caccattaaa	gctgcttacc	aagaacttcc	ccagcatttt	gacttccttg	tttgatagct	60
------------	------------	------------	------------	------------	------------	----

```
<210> 68
<211> 551
<212> DNA
<213> Homo sapien
```

<400>	68						
tagct	ggtacataat	cactgaggag	ctattttctta	acatgctttt	atagaccatg		60
gctag	accagtat	aagggtaat	ctcacacctc	cttagctgta	agagtctggc		120
acaga	cctctctgtg	caataacttg	tggccactgg	aatccctgg	gccggcattt		180
gggg	tgcaatgact	ccaagggcc	aaaagagtta	aaggcacgac	tgggatttct		240
gactg	tggtgaaact	ccttccaagg	ctgaggggg	cagtangtgc	tctgggaggg		300
gcacc	actttgat	tcaacaagcc	acttgaagcc	caattataaa	attgttattt		360
ctgat	ggaactcaat	ttgaaccttc	aaaactttgt	tagtttatcc	tattatattg		420
cctaa	ttacatttgt	ctagcattgg	atttggttcc	tgtngcatat	gtttttttcn		480
tggtgc	cccccccc	nnatctta	ttaaaccnca	attttgcnat	tcnccnnnnn		540
annna	a						551

```
<220>
<221> misc_feature
<222> (1)...(396)
<223> n = A,T,C or G
```

$\langle 210 \rangle$	70
$\langle 211 \rangle$	536

<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(536)
<223> n = A,T,C or G

<400> 70

actagtgcaa	aagcaaatat	aaacatcgaa	aaggcggttc	tcacgttagc	tgaagatata	60
cttcgaaaga	cccctgtaaa	agagcccaac	agtgaaaatg	tagatatcag	cagtggagga	120
ggcgtgacag	gctggaagag	caaatgctgc	tgagcattct	cctgttccat	cagttgccat	180
ccactacccc	gttttctctt	cttgctgcaa	aataaaccac	tctgtccatt	tttaactcta	240
aacagatatt	tttgtttctc	atcttaacta	tccaagccac	ctattttatt	tgttctttca	300
tctgtgactg	cttgctgact	ttatcataat	tttcttcaaa	caaaaaaatg	tatagaaaaa	360
tcatgtctgt	gacttcattt	ttaaatgnta	cttgctcagc	tcaactgcat	ttcagttggt	420
ttatagtcca	gttcttatca	acattnaaac	ctatngcaat	catttcaaat	ctattctgca	480
aattgtataa	gaataaaagt	tagaatttaa	caattaaaaa	aaaaaaaaaa	aaaaaa	536

<210> 71
<211> 865
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(865)
<223> n = A,T,C or G

<400> 71

gacaaagcgt	taggagaaga	anagaggcag	ggaanaactnc	ccaggcacga	tggccncctt	60
cccaccagca	accagcggcc	cccaccagcc	cccaggcccg	gacgacgaag	actccatcct	120
ggattaatct	nacctctntc	gcctgnccca	ttcttacctc	ggaggtggag	gccggaaagg	180
tcncaccaag	aganaaactg	ctgccaaac	caaccgcccc	agccctggcg	ggcacganag	240
gaaactggtg	accaatctgc	agaattctna	gaggaanaag	cnagggggccc	cgcgctnaga	300
cagagctgga	tatgangcca	gaccatggac	netacnccn	ncaatncana	cgggactgcg	360
gaagatggan	gacccnccgac	nngatcagge	cngetnncca	nccccccacc	cctatgaatt	420
attcccgcgtg	aangaatctc	tgannggctt	ccannaaagc	gcctccccnc	cnaacgnaan	480
tncaacatng	ggattanang	ctggggaactg	naaggggcaa	ancctnnaat	atccccagaa	540
acaanctctc	ccnaanaaac	tggggcncct	catnggtggn	accaactatt	aactaaaccg	600
cacgccaagn	aantataaaa	ggggggcccc	tcnccggng	accccccttt	gtcccttaat	660
ganggttata	cnccttgctg	accatggtnc	ccnnttctgt	ntgnatgttt	ccnctcccct	720
ccnctatnt	cnagccgaac	tcnnatttnc	cgggggggtgc	natchnantng	tnncctttt	780
ttngttgncc	cngccctttc	cngcggaaacn	cgtttccccg	ttantaacgg	caccgggggn	840
aagggtgntt	ggccccctcc	ctccc				865

<210> 72
<211> 560
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(560)

<223> n = A,T,C or G

<400> 72

```
cctggacttg tcttggttcc agaacctgac gaccgcgga cggcgacgtc tcttttgaact 60
aaaagacagt gtccagtgt cngcctagg agtctacggg gaccgcctcc cgcgcgcgcca 120
ccatgcccaa cttctctggc aactggaaaa tcatccgac ggaaaacttc gangaattgc 180
tcnaantgct gggggtgaat gtgatgctna ngaanattgc tgtggctgca gcgtccaagc 240
cagcagtgga gatcnaacag gagggagaca ctttctacat caaaacctcc accaccgtgc 300
gcaccacaaa gattaacttc nnngttgggg aggantttga ggancaaaact gtggatngga 360
ngcctgtnaa aacctggtga aatgggagaa tganaataaa atggtctgtg ancanaaaact 420
cctgaaagga gaaggccccc anaactcctg gaccngaaaa actgaccnc cnatngggga 480
actgatnctt gaacctgaa cgggcgggat ganccttttt tnttgcncnc naanggggtc 540
tttcnntttc cccaaaaaaa
```

<210> 73

<211> 379

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(379)

<223> n = A,T,C or G

<400> 73

```
ctggggancc ggcggtnngc nccatntcnn gncgcgaagg tggcaataaa aancnctga 60
aaccgcncaa naacatgcc naagatatgg acgaggaaga tngngctttc nngnacaanc 120
gnanngagga acanaacaaa ctenangagc tctcaagcta atgccgcggg gaagggggccc 180
ttggccacnn gtggaattaa gaaatctggc aaanngtann tgttccttgt gcctnangag 240
ataaangacc ctttattttca tctgtattta aacctctctn ttccctgnca taacttcttt 300
tnccacgtan agntggaant anttggtgtc ttggactgtt gtncatttta gannaaaactt 360
ttgttcaaaa aaaaaataa
```

<210> 74

<211> 437

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(437)

<223> n = A,T,C or G

<400> 74

```
actagttcag actgccacgc caaccccaga aaatacccca catgccagaa aagtgaagtc 60
ctaggtgttt ccatctatgt ttcaatctgt ccatctacca ggctctcgca taaaaacaaa 120
acaaaaaaaa gctgccaggt tttanaagca gttctggtct caaaaccatc aggatcctgc 180
caccagggtt cttttgaaat agtaccacat gtaaaaggga atttggcttt cacttcatct 240
aatcactgaa ttgtcaggct ttgattgata attgtagaaa taagtagcct tctgttgtgg 300
gaataagtta taatcagtat tcatctcttt gttttttgtc actcttttct ctctnattgt 360
gtcatttgta ctgtttgaaa aatatttctt ctataaaatt aaactaacct gccttaaaaa 420
aaaaaaaaaa aaaaaaa
```

<210> 75

<211> 579
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(579)
 <223> n = A,T,C or G

<400> 75

ctccgctcgcc	gccaaagatga	tgtgctggggc	gccctccgcc	acgcagccgg	ccaccgccga	60
gacccagcac	atcgccgacc	aggtgaggtc	ccagcttgaa	gagaaagaaa	acaagaagtt	120
ccctgtgttt	aaggccgtgt	cattcaagag	ccaggtggc	gcggggacaa	actacttcat	180
caaggtgcac	gtcggcgacg	aggacttcgt	acacctgcga	gtgttccaat	ctctccctca	240
tgaaaacaag	cccttgacct	tatctaacta	ccagaccaac	aaagccaagc	atgatgagct	300
gacctatttc	tgatcctgac	tttggacaag	gcccttcagc	cagaagactg	acaaagtcac	360
cctccgtcta	ccagagcgtg	cacttgtgat	cctaaaataa	gcttcacctc	cgggctgtgc	420
ccttgggggtg	gaagggggcan	gatctgcact	gcttttgcac	ttctcttcc	aaatttcatt	480
gtgttgattc	tttccttcca	ataggtgatc	ttnattactt	tcagaatatt	ttccaaatna	540
gatataattt	naaaatcctt	aaaaaaaaa	aaaaaaaaa			579

<210> 76
 <211> 666
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(666)
 <223> n = A,T,C or G

<400> 76

gtttatccta	tctctccaac	cagattgtca	gtccttgag	ggcaagagcc	acagtatatt	60
tccctgtttc	ttccacagt	cctaataata	ctgtggaact	aggttttaat	aatttttta	120
ttgatgttgt	tatggcgagg	atggcaacca	gaccattgtc	tcagagcagg	tgctggctct	180
ttcctggcta	ctccatgttg	gctagcctct	ggtaacctct	tacttattat	cttcaggaca	240
ctcactacag	ggaccaggga	tgatgcaaca	tccttgtctt	tttatgacag	gatgtttgct	300
cagcttctcc	aacaataaaa	agcacgtggg	aaaacacttg	cggatattct	ggactgtttt	360
taaaaaatat	acagtttacc	gaaaatcata	ttatcttaca	atgaaaagga	ntttatagat	420
cagccagtga	acaacctttt	cccaccatac	aaaaattcct	tttcccgaan	gaaaangget	480
ttctcaataa	ncctcacttt	cttaanatct	tacaagatag	ccccganatc	ttatcgaaac	540
tcatttttagg	caaatatgan	ttttattgtg	cgttacttgt	ttcaaaattt	ggtattgtga	600
atatcaatta	ccacccccat	ctcccatgaa	anaaanggga	aanggtgaan	ttcntaancg	660
cttaaa						666

<210> 77
 <211> 396
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(396)
 <223> n = A,T,C or G

<400> 77

ctgcagcccg	ggggatccac	taatctacca	nggttatttg	gcagctaatt	ctanatttgg	60
atcattgccc	aaagttgcac	ttgctgggtct	cttgggattt	ggccttgga	aggtatcata	120
catanganta	tgccanaata	aattccattt	ttttgaaaat	canctccttg	gggctggttt	180
tggtccacag	cataacangc	actgcctcct	tacctgtgag	gaatgcaaaa	taaagcatgg	240
attaagtgag	aagggagact	ctcagccttc	agcttccctaa	attctgtgtc	tgtgactttc	300
gaagtttttt	aaacctctga	atttgtacac	attttaaatt	tcaagtgtac	ttttaaataa	360
aatacttcta	atgggaacaa	aaaaaaaaaa	aaaaaa			396

<210> 78

<211> 793

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(793)

<223> n = A,T,C or G

<400> 78

gcacccctagc	cgccgactca	cacaaggcag	gtgggtgagg	aaatccagag	ttgccatgga	60
gaaaattcca	gtgtcagcat	tcttgcctct	tgtggccctc	tcctacactc	tgccagaga	120
taccacagtc	aaacctggag	ccaaaaagga	cacaaaggac	tctcgaccca	aactgcccc	180
gacctctctc	agaggttggg	gtgaccaact	catctggact	cagacatatg	aagaagctct	240
atataaatcc	aagacaagca	acaaaccctt	gatgattatt	catcacttgg	atgagtgcc	300
acacagtcna	gctttaaaga	aagtgtttgc	tgaaaataaa	gaaatccaga	aattggcaga	360
gcagtttgtc	ctcctcaatc	tggtttatga	aacaactgac	aaacaccttt	ctcctgatgg	420
ccagtatgtc	ccaggattat	gtttgttgac	ccatctctga	cagttgaagc	cgatatcctg	480
ggaagatatt	cnaaccgtct	ctatgcttac	aaactgcaga	tacgctctgt	tgcttgacac	540
atgaaaaagc	tctcaagttg	ctnaaaatga	attgtaagaa	aaaaaatctc	cagccttctg	600
tctgtcggct	tgaaaattga	aaccagaaaa	atgtgaaaaa	tggtatttgt	ggaacanatn	660
gacacctgat	taggttttgg	ttatgttcac	cactattttt	aanaaaanan	nttttaaaat	720
ttggttcaat	tntctttttn	aaacaatntg	tttctacntt	ngnancgtgat	ttctaaaaaa	780
aataatnttt	ggc					793

<210> 79

<211> 456

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(456)

<223> n = A,T,C or G

<400> 79

actagtatgg	ggtgggaggc	cccacccttc	tcccctaggc	gctgtttctg	ctccaaaggg	60
ctccgtggag	agggactggc	agagctgang	ccacctgggg	ctggggatcc	cactcttctt	120
gcagctgttg	agcgcaccta	accactggtc	atgccccac	ccctgctctc	cgcacccgct	180
tcctcccgac	cccangacca	ggctacttct	cccctcctct	tgcctccctc	ctgcccctgc	240
tgcctctgat	cgtangaatt	gangantgtc	cgccttgtgt	gctganaatg	gacagtggca	300
ggggctggaa	atgggtgtgt	gtgtgtgtgt	gtgtgtgtgt	gtgtgtgtgt	gcnccccccc	360
tgaagaccg	agattgaggg	aaancatgtc	tgtctgggtgt	gacctgtttt	cctctccata	420

456

```
<210> 80
<211> 284
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(284)
<223> n = A,T,C or G
```

<400> 80

ctttgtacct	ctagaaaaga	taggtattgt	gtcatgaaac	ttgagtttaa	attttatata	60
taaaactaaa	agtaatgctc	acttttagcaa	cacatactaa	aattggaacc	atactgagaa	120
gaatagcatg	acctccgtgc	aaacaggaga	agcaaatttg	tgtatgtgtg	attaaaaaga	180
aataaataaa	tgtgtatatg	tgtaaacttg	atgtttattg	ggaatacaga	ttgggaaata	240
aaatgtattt	cttactgtga	aaaaaaaaaa	aaaaaaaaaa	aana		284

```
<210> 81
<211> 671
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(671)
<223> n = A,T,C or G
```

<400> 81

gccaccaaca	ttccaagcta	ccctgggtac	ctttgtgcag	tagaagctag	tgagcatgtg	60
agcaagcggg	gtgcacacgg	agactcatcg	ttataattta	ctatctgcc	agagtagaaa	120
gaaaggctgg	ggatatattg	gttggcttgg	ttttgatttt	ttgcttgttt	gtttgttttg	180
tactaaaaca	gtattatctt	ttgaatatcg	tagggacata	agtatataca	tgttatccaa	240
tcaagatggc	tagaatgggtg	cctttctgag	tgtctaaaac	ttgacacccc	tggtaaatct	300
ttcaacacac	ttccactgcc	tgcgtaatga	agttttgatt	catttttaac	cactggaatt	360
tttcaatgcc	gtcatttttca	gttagatnat	tttgcacttt	gagattaaaa	tgccatgtct	420
atttgattag	tcttattttt	ttattttttac	aggcttatca	gtctcactgt	tggtgtcat	480
tgtgacaaa	g	g	g	g	g	540
acattaaagt	ttggccaaaa	aatgttgc	gtgttttacc	tcgacttgct	aatcaatan	600
canaaaggct	ggctnataat	g	g	g	g	660
aaaaaaaaa	a	g	g	g	g	671

```
<210> 82
<211> 217
<212> DNA
<213> Homo sapien
```

```
<220>
<221> misc_feature
<222> (1)...(217)
<223> n = A,T,C or G
```

<400> 82

```
<210> 83
<211> 460
<212> DNA
<213> Homo sapien
```

<400> 83

```
<210> 84
<211> 323
<212> DNA
<213> Homo sapien
```

<400> 84

```
<210> 85
<211> 771
<212> DNA
<213> Homo sapien
```

<400> 85

```
<210> 86
<211> 628
<212> DNA
<213> Homo sapien
```

<400> 86

```
<210> 87
<211> 518
<212> DNA
<213> Homo sapien
```

<400> 87

ttttttattt	tttttagaga	gtagttcagc	ttttatttat	aaatttattg	cctgttttat	60
tataacaaca	ttatactggt	tatggtttaa	tacatatggt	tcaaaatgta	taatacatca	120
agtagtacag	ttttaaaatt	ttatgcttaa	aacaagtttt	gtgtaaaaaa	tgcagataca	180
ttttacatgg	caaatcaatt	tttaagtcac	cctaaaaatt	gatttttttt	tgaaatttaa	240
aaacacattt	aattttcaatt	tctctcttat	ataaccttta	ttactatagc	atgggtttcca	300
ctacagttta	acaatgcagc	aaaattccca	tttcacggta	aattggggtt	taagcggcaa	360

```

ggttaaaatg ctttgaggat cctnaatacc ctttgaactt caaatgaagg ttatggttgt      420
naatttaacc ctcatgccat aagcagaagc acaagtttag ctgcattttg ctctaaactg      480
taaaancgag cccccgttg aaaaagcaaa agggaccc                                518

```

```

<210> 88
<211> 1844
<212> DNA
<213> Homo sapien

```

```
<400> 88
```

```

gagacagtga atcctagtat caaaggattt ttggcctcag aaaaagttgt tgattatttt      60
tattttattt tatttttcga gactccgtct caaaaaaaaa aaaaaaaaaa agaatcacaa      120
ggtattttgct aaagcatttt gagctgcttg gaaaaaggga agtagttgca gtagagtttc      180
ttccatcttc ttggtgcttg gaagccatat atgtgtcttt tactcaagct aaggggtata      240
agcttatgtg ttgaatttgc tacatctata ttccacatat tctcacata agagaatttt      300
gaaatagaaa tatcatagaa catttaagaa agtttagtat aaataatatt ttgtgtgttt      360
taatcccttt gaagggatct atccaaagaa aatattttac actgagctcc ttcctacacg      420
tctcagtaac agatcctgtg ttagtctttg aaaatagctc atttttttaa tgtcagttag      480
tagatgtagc atacatatga tgtataatga cgtgtattat gttaacaatg tctgcagatt      540
ttgtaggaat acaaaacatg gcctttttta taagcaaaac gggccaatga ctagaataac      600
acatagggca atctgtgaat atgtattata agcagcattc cagaaaagta gttggtgaaa      660
taattttcaa gtcaaaaagg gatatggaaa ggaattatg agtaacctct attttttaag      720
ccttgctttt aaattaaacg ctacagccat ttaagccttg aggataataa agcttgagag      780
taataatgtt aggttagcaa aggttttagat gtatcacttc atgcatgcta ccatgatagt      840
aatgcagctc ttcgagtcac ttctggctcat tcaagatatt cacccttttg cccatagaaa      900
gcaccctacc tcacctgctt actgacattg tcttagctga tcacaagatc attatcagcc      960
tccattattc cttactgtat ataaaataca gagttttata ttttcctttc ttogtttttc     1020
accatattca aaacctaaat ttgttttttg agatggaatg caaagtaatc aagtgttcgt     1080
gctttcacct agaagggtgt ggctcctgaag gaaagaggtc cctaaatatc cccaccctg     1140
ggtgctcctc cttccctggg accctgacta ccagaagtcg ggtgctagag cagctggaga     1200
agtgcagcag cctgtgcttc cacagatggg ggtgctgctg caacaaggct ttcaatgtgc     1260
ccatcttagg gggagaagct agatcctgtg cagcagcctg gtaagtectg aggaggttcc     1320
attgctcttc ctgctgctgt cctttgcttc tcaacggggc tcgctctaca gtctagagca     1380
catgcagcta acttgtgcct ctgcttatgc atgagggtta aattaacaac cataaccttc     1440
atttgaagtt caaagggtga ttcaggatcc tcaaagcatt ttaaccttgc cgcttaaaac     1500
ccaatttacc gtgaaatggg aattttgctg cattgtttaa ctgtagtgga aaccatgcta     1560
tagtaataaa gggtatataa gagagaaatt gaaattaaat gtgtttttta atttcaaaaa     1620
aaaatcaatc tttaggatga cttaaaaaatt gatttgccat gtaaaatgta tctgcatttt     1680
ttacacaaaa cttgttttta gcataaaatt ttaaaactgt actacttgat gtattatata     1740
ttttgaacca tatgtattaa accataaaca gtataatgtt gttataataa aacaggcaat     1800
aaatttataa ataaaagctg aaaaaaaaaa aaaaaaaaaa aaaa                                1844

```

```

<210> 89
<211> 523
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(523)
<223> n = A,T,C or G

```

```
<400> 89
```

```

tttttttttt tttttttagt caatccacat ttattgatca cttattatgt accaggcact      60

```

```

gggataaaga tgactgttag tcaactcacag taaggaagaa aactagcaaa taagacgatt 120
acaatatgat gtagaaaatg ctaagccaga gatatagaaa ggtcctattg ggtccttctg 180
tcaccttgtc tttccacatc cctacccttc acaggccttc cctccagctt cctgcccccg 240
ctccccactg cagatcccct gggattttgc cttagagctaa acgagganat gggccccctg 300
gccctggcat gacttgaacc caaccacaga ctgggaaagg gagcctttcg anagtggatc 360
actttgatna gaaaacacat aggggaattga agagaaantc cccaaatggc caccctgtgt 420
ggtgctcaag aaaagtgtgc agaattggata aatgaaggat caagggaatt aatanatgaa 480
taattgaatg gtggctcaat aagaatgact ncnttgaatg acc 523

```

```

<210> 90
<211> 604
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(604)
<223> n = A,T,C or G

```

```

<400> 90
ccagtgtggt ggaatgcaaa gattacccccg gaagcctttcg agaagctggg attccctgca 60
gcaaaggaaa tagccaatat gtgtcggtttc tatgaaatga agccagaccg agatgtcaat 120
ctcaccacc aactaaatcc caaagtcaaa agcttcagcc agtttatctc agagaaccag 180
gggagccttc aagggcattg agaaaatcag ctgttcagat aggcctctgc accacacagc 240
ctctttcctc tctgatcctt ttctctctta cggcacaaca ttcatgtttg acagaacatg 300
ctggaatgca attgtttgca acaccgaagg atttctctgc gtgcctctt cagtaggaag 360
cactgcattg gtgataggac acggtaattt gattcacatt taacttgcta gttagtata 420
aggggtggtgta cactgtttt gtaaaatgag aagcctcgga aacttgggag cttctctcct 480
accactaatg gggaggggcag attattactg ggattttctc tggggtgaat taatttcaag 540
ccctaattgc tgaaattccc ctnggcaggc tccagttttc tcaactgcat tgcaaaattc 600
cccc 604

```

```

<210> 91
<211> 858
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(858)
<223> n = A,T,C or G

```

```

<400> 91
tttttttttt ttttttttta tgattattat tttttttatt gatctttaca tcttcagtgt 60
tggcagagtt tctgatgctt aataaacatt tgttctgata agataagtgg aaaaaattgt 120
catttcctta ttcaagccat gctttttctgt gatattctga tcttagttga acatacagaa 180
ataaatgtct aaaacagcac ctcgattctc gtctataaca ggactaagtt cactgtgata 240
ttaaataagc ttggctaaaa tgggacatga gtggaggtag tcacacttca gcgaagaaag 300
agaatctcct gtataatctc accaggagat tcaacgaatt ccaccacact ggactagtgg 360
atcccccggt ctgcaggaaat tcgatatcaa gcttatcgat accgtcgacc tcgagggggg 420
gcccggtacc caattcgccc tatagttagt cgtattacgc gcgctcactg gccgtcgttt 480
tacaacgtcg tgactgggaa aacctgggcg ttaccaact taatcgctt gcagcacatc 540
ccctttctgc cagctggcgt aatagcgan agccgcacc gatcgccctt ncaacagttg 600
cgcagcctga atggcgaatg ggacgcgccc tgtagcggcg cattaaagcg cggcnggggtg 660

```

tggnggntcc	cccacgtgac	cgntacactt	ggcagcgcct	tacgccggtc	nttcgctttc	720
ttcccttcct	ttctcgcacc	gttcgcgggg	tttccccgnn	agctnttaat	cgggggncctc	780
cctttanggg	tncnaattaa	nggnttacong	gaccttngan	cccaaaaact	ttgattaggg	840
ggaaggtccc	cgaagggg					858

<210> 92
 <211> 585
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(585)
 <223> n = A,T,C or G

<400> 92						
gttgaatctc	ctggtgagat	tatacaggag	attctctttc	ttcgctgaag	tgtgactacc	60
tccactcatg	tcccatttta	gccaaagctta	tttaagatca	cagtgaactt	agtcctgtta	120
tagacgagaa	tcgaggtgct	gttttagaca	tttatttctg	tatgttcaac	taggatcaga	180
atatcacaga	aaagcatggc	ttgaataagg	aaatgacaat	tttttccact	tatctgatca	240
gaacaaatgt	ttattaagca	tcagaaactc	tgccaacact	gaggatgtaa	agatcaataa	300
aaaaaataat	aatcatnann	naaanannan	nngaagggcg	gccgccaccg	cgggtggagct	360
ccagcttttg	ttcccttttag	tgagggttaa	ttgcgcgctt	ggcggttaatc	atggtcatag	420
ctgtttcctg	tgtgaaattg	ttatccggct	cacaattccn	cncaacatac	gagccgggaa	480
gcntnangtg	taaaagcctg	gggggtgcta	attgagtgag	ctnactcaca	ttaattgngt	540
tgcgctccac	ttgcccgcctt	ttccantcog	ggaaacctgt	tcgnc		585

<210> 93
 <211> 567
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(567)
 <223> n = A,T,C or G

<400> 93						
cggcagtgtt	gctgtctgcg	tgtccacott	ggaatctggc	tgaactggct	gggaggacca	60
agactgcggc	tgggggtggc	anggaaggga	accgggggct	gctgtgaagg	atcttggaac	120
ttccctgtac	ccaccttccc	cttgcctcat	gtttgtanag	gaaccttgtg	ccggccaagc	180
ccagtttcct	tgtgtgatac	actaatgtat	ttgctttttt	tgggaaatan	anaaaaaatca	240
attaaattgc	tantgtttct	ttgaannnnn	nnnnnnnnnn	nnnnnnnggg	gggngcgccc	300
ccnccgngga	aacnccccct	tttgttccct	ttaattgaaa	ggttaattng	cncnctggc	360
gttaancctt	gggccaaaanc	tngttncccg	tgntgaaatt	gttnatcccc	tcccaaattc	420
ccccccncc	ttccaaaccc	ggaaancctn	annntgttna	ancccggggg	gttgccctaan	480
ngnaattnaa	ccnaaccccc	ntttaaatng	nnnttgncn	ccacnngccc	cncctttccca	540
nttcggggaa	aacctnttcc	gtgccca				567

<210> 94
 <211> 620
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(620)
 <223> n = A,T,C or G

<400> 94

actagtcaaa	aatgctaaaa	taatttggga	gaaaatattt	tttaagtagt	gttatagttt	60
catgtttatc	ttttattatg	ttttgtgaag	ttgtgtcttt	tcactaatta	cctatactat	120
gccaatattt	ccttatatct	atccataaca	tttatactac	atttgtaana	naatatgcac	180
gtgaaactta	acactttata	aggtaaaaat	gaggtttcca	anatttaata	atctgatcaa	240
gttcttggta	tttccaaata	gaatggactt	ggtctgttaa	gggctaagga	gaagaggaag	300
ataagggtta	aagttgttaa	tgaccaaaca	ttctaaaaga	aatgcaaaaa	aaaagtttat	360
tttcaagcct	tcgaactatt	taaggaaagc	aaaatcattt	cctaaatgca	tatcatttgt	420
gagaatttct	cattaatatc	ctgaatcatt	catttcacta	aggctcatgt	tnactccgat	480
atgtctctaa	gaaagtacta	tttcatggtc	caaacctggt	tgccatantt	gggttaaaggc	540
tttcccttaa	gtgtgaaant	atttaaaatg	aaattttcct	ctttttaaaa	attctttana	600
agggtttaagg	gtgttgggga					620

<210> 95
 <211> 470
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(470)
 <223> n = A,T,C or G

<400> 95

ctcgaccttc	tctgcacagc	ggatgaaccc	tgagcagctg	aagaccagaa	aagccactat	60
nactttntgc	ttaattcang	agcttacang	attcttcaaa	gagtgngtcc	agcatccttt	120
gaaacatgag	ttcttaccag	cagaagcaga	cctttacccc	accacctcag	cttcaacagc	180
agcaggtgaa	acaacccatc	cagcctccac	ctnaggaaat	atttgttccc	acaaccaagg	240
agccatgcca	ctcaaagggt	ccacaacctg	naaacacaaa	nattccagag	ccaggctgta	300
ccaaggtccc	tgagccaggg	ctgtaccaan	gtccctgagc	caggttgtag	caangtccct	360
gagccaggat	gtaccaagg	cctgancca	ggttggtccaa	ggtccctgag	ccagggtaca	420
ccaagggcct	gngccaggca	gcatacaangt	ccttgacca	ggcttatcaa		470

<210> 96
 <211> 660
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(660)
 <223> n = A,T,C or G

<400> 96

tttttttttt	tttttttttt	ggaattaaaa	gcaatttaat	gagggcagag	caggaaacat	60
gcatttcttt	tcattcgaat	cttcagatga	accctgagca	gccgaagacc	agaaaagcca	120
tgaagacttt	ctgcttaatt	caggggctta	caggattctt	cagagtgtgt	gtgaacaaaa	180
gctttatagt	acgtattttt	aggatacaaa	taagagagag	actatggctt	gggggtgagaa	240
tgtactgatt	acaaggctta	cagacaatta	agacacagaa	acagatggga	agagggtgnc	300

```
<210> 97
<211> 441
<212> DNA
<213> Homo sapien
```

<400>		97							
catac	anagtattcc	tctettcacaca	ccaggaccag	ccactgttgc	agcatgagtt			60	
cagca	gaagcagccc	tgcacccccac	cccctcagct	tcagcagcag	caggtgaaac			120	
cgcca	gcctocacct	caggaaccat	gcacccccaa	aaccaaggag	ccctgccacc			180	
gtgcc	tgagccctgc	caccccaaag	tgccctgagcc	ctgccagccc	aaggttccag			240	
tgcca	ccccaaagtg	cctgagccct	gccctttaat	agtcactcca	gcaccagccc			300	
aanac	caagcagaag	taatgtggto	cacagccatg	cccttgagga	gccggccacc			360	
ctgae	tcccttatcc	cattcttgtt	atgagtccca	tttgccctgc	aatttagcatt			420	
tcccc	aaaaaaaaaa	a						441	

```
<210> 98
<211> 600
<212> DNA
<213> Homo sapien
```

```
<220>
<221> misc_feature
<222> (1)...(600)
<223> n = A,T,C or G
```

```
<210> 99
<211> 667
<212> DNA
<213> Homo sapien
```

<220>
 <221> misc_feature
 <222> (1)...(667)
 <223> n = A,T,C or G

<400> 99

actagtgact	gagttcctgg	caaagaaatt	tgacctggac	cagttgataa	ctcatgtttt	60
accattttaa	aaaatcagtg	aaggatttga	gctgctcaat	tcaggacaaa	gcattcgaac	120
ggtcctgacg	ttttgagatc	caaagtggca	ggaggtctgt	gttgctcatg	tgaactggag	180
tttctcttgt	gagagttccc	tcactctgaa	tcagtgtatc	gtctcacaaa	tacaagcata	240
agtagaagat	ttgttgaaag	catagaaccc	ttataaagaa	ttattaacct	ttataaacat	300
ttaaagtctt	gtgagcacct	gggaattagt	ataataacaa	tgttnatatt	tttgatttac	360
atthttgtaag	gctataattg	tatcttttaa	gaaaacatac	cttggaatttc	tatgttgaaa	420
tggagatttt	taagagtttt	aaccagctgc	tgcagatata	ttactcaaaa	cagatatagc	480
gtataaagat	atagtaaatg	catctcctag	agtaatatcc	acttaacaca	ttggaaacta	540
ttatttttta	gatttgaaata	tnaatgttat	tttttaacaa	cttggttatga	gttacttggg	600
attacatttt	gaaatcagtt	cattccatga	tgcantattac	tgggattaga	ttaagaaaga	660
cggaaaaa						667

<210> 100
 <211> 583
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(583)
 <223> n = A,T,C or G

<400> 100

gttttggttg	taagatgatc	acagtcattg	tacactgate	taaaggacat	atatataacc	60
cttttaaaaa	aaaatcactg	cctcattctt	atttcaagat	gaatttctat	acagactaga	120
tgtttttctg	aagatcaatt	agacattttg	aaaatgattt	aaagtgtttt	ccttaattgtt	180
ctctgaaaac	aagtttcttt	tgtagtttta	acaaaaaaag	tgcccttttt	gtcactggat	240
tctcctagca	ttcatgatth	ttttttcata	caatgaaatt	aaaattgcta	aaatcatgga	300
ctggctttct	ggttggaatt	caggtaagat	gtgtttaagg	ccagagcttt	tctcagtatt	360
tgattttttt	ccccaatatt	tgatttttta	aaaatataca	catnggtgct	gcattttatat	420
ctgctgggtt	aaaattctgt	catattttcac	ttctagcctt	ttagttatgg	caaatcatat	480
tttactttta	cttaaaagcat	ttggtnattt	ggantatctg	gttctannct	aaaaaaaanta	540
attctatnaa	ttgaantttt	ggtactcnnc	catatttgga	tcc		583

<210> 101
 <211> 592
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(592)
 <223> n = A,T,C or G

<400> 101

gtggagacgt	acaaagagca	gccgctcaag	acacctggga	agaaaaagaa	aggcaagccc	60
gggaaacgca	aggagcagga	aaagaaaaaa	cggcgaactc	gctctgcctg	gttagactct	120

```

ggagtgactg ggagtgggct agaaggggac cacctgtctg acacctccac aacgtcgtctg 180
gagctcgatt cacggaggca ttgaaatddd cagcaganac cttccaagga catattgcag 240
gattctgtaa tagtgaacat atggaaagta ttagaaatat ttattgtctg taaataactgt 300
aaatgcattg gaataaaaact gtctccccc a ttgctctatg aaactgcaca ttgggtcattg 360
tgaatatttt tttttttgcc aaggctaata caattattat tatcacattt accataattt 420
atdddgtcca ttgatgtatt tatdddgtaa atgtatcttg gtgctgctga attdctatat 480
ttdddgtaca taatgcnttt anatatacct atcaagtttg ttgataaatg acncaatgaa 540
gtgncncnan ttgngngttg aattdaatga atgcctaatt ttattatccc aa 592

```

```

<210> 102
<211> 587
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(587)
<223> n = A,T,C or G

```

```

<400> 102
cgtcctaagc acttagacta catcaggga gaacacagac cacatccctg tcctcatgctg 60
gcttatgttt tctggaagaa agtggagacc nagtccttgg ctttagggct ccccggtctg 120
gggctgtgca ntccggtcag ggcgggaagg gaaatgcacc gctgcatgtg aacttacagc 180
ccaggcggat gccccttccc ttagcactac ctggcctcct gcacccctc gcctcatgtt 240
cctcccacct tcaanaaatg aanaacccca tgggccccagc cccttgccct ggggaaccaa 300
ggcagccttc caaaactcag gggctgaagc anactattag ggcaggggct gactttgggt 360
gacactgcc attccctctc agggcagctc angtcaccn ggnetcttga acccagcctg 420
ttcctttgaa aaagggcaaa actgaaaagg gcttttccta naaaaagaaa aaccagggaa 480
ctttgccagg gcttcnntnt taccaaaaacn ncttctcnng gatttttaat tccccattng 540
gcctccactt accnggggcn atgccccaaa attaanaatt tcccatc 587

```

```

<210> 103
<211> 496
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(496)
<223> n = A,T,C or G

```

```

<400> 103
anaggactgg ccctacntgc tctctctcgt cctacctatc aatgcccac atggcagaa 60
ctgcanccct tggncactgc anatggaaac ctctcagtgt cttgacatca ccctaccnt 120
gcggtgggtc tccaccacaa ccactttgac tctgtgggtc ctgnanggtg gnttctcctg 180
actggcagga tggacottan ccnacatata cctctgttcc ctctgctnag anaaagaatt 240
cccttaacat gatataatcc acccatgcaa ntngctactg gccagctac catttaccat 300
ttgcctacag aatttcattc agtctacaact ttggcattct ctctggcgat agagtgtggc 360
tgggctgacc gcaaaagggtg ccttacacac tggccccac cctcaaccgt tgacncatca 420
gangcttgcc tcctccttct gattnncccc catgttggat atcaggggtg tcnagggatt 480
ggaaaagaaa caaaac

```

```

<210> 104
<211> 575

```

<212> DNA
 <213> Homo sapien
 <220>
 <221> misc_feature
 <222> (1)...(575)
 <223> n = A,T,C or G

<400> 104

```
gcacctgctc tcaatccnnc tctcaccatg atcctccgcc tgcanaaact cctctgccaa      60
ctatggangt ggtttcnggg gtggctcttg ccaactggga agaagccgtg gtgtctctac      120
ctgttcaact cngtttgtgt ctgggggata aactnggggc tatggaagcg gctnaactgt      180
tgttttggtg gaagggctgg taattggott tgggaagtng cttatngaag ttggcctngg      240
gaagttgcta ttgaaaagtng ccntggaagt ngntttggtg gggggttttg ctggtggcct      300
ttgttnaatt tgggtgcttt gtnaatggcg gccccctcnc ctgggcaatg aaaaaaatca      360
ccnatgcngn aaacctcnac nnaacagcct gggcttcctc cacctcgaaa aaagttgctc      420
ccccccaaa aaaggncaan cccctcaann tgggaangttg aaaaaatcct cgaatgggga      480
nccnnaaac aaaaancccc ccntttcccn gnaanggggg aaataccncc ccccactta      540
cnaaaacct tntaaaaaac ccccgggaa aaaaa
```

<210> 105
 <211> 619
 <212> DNA
 <213> Homo sapien
 <220>
 <221> misc_feature
 <222> (1)...(619)
 <223> n = A,T,C or G

<400> 105

```
cactagtagg atagaaacac tgtgtccga gagtaaggag agaagctact attgattaga      60
gcctaaccga ggttaactgc aagaagaggc gggatacttt cagctttcca tgtaactgta      120
tgcataaagc caatgtagtc cagtttctaa gatcatgttc caagctaaact gaatcccact      180
tcaatacaca ctcatgaact cctgatggaa caataacagg cccaagcctg tggtatgatg      240
tgcacacttg ctgactcan aaaaaatact actctcataa atgggtggga gtattttggt      300
gacaacctac tttgcttggc tgagtgaagg aatgatattc atatattcat ttattccatg      360
gacatttagt tagtgctttt tatataccag gcatgatgct gagtgacact cttgtgtata      420
tttccaaatt tttgtacagt cgctgcacat atttgaaatc atatattaag acttccaaaa      480
aatgaagtcc ctggtttttc atggcaactt gatcagtaaa ggattcnctc ctggttggtg      540
cttaaaacat ctactatatn gttanatga aattcctttt ccccnctcc cgaaaaaana      600
aagtgggtgg gaaaaaaaa
```

<210> 106
 <211> 506
 <212> DNA
 <213> Homo sapien
 <220>
 <221> misc_feature
 <222> (1)...(506)
 <223> n = A,T,C or G

<400> 106

```
<210> 107
<211> 452
<212> DNA
<213> Homo sapien
```

<400>	107						
gtctg	tactaaacag	taagatatct	caatgaacca	taaattcaac	tttgtaaaaa		60
tgaag	catagataat	attgtttggg	aaatgtttct	tttgtttggg	aaatgtttct		120
agacc	ctcctattct	ataaaactct	gcatgtagag	gcttgtttac	ctttctctct		180
gttta	caataggagt	ggtgatttga	aaaatataaa	attatgagat	tgggttttct		240
ataaa	ttgcatcact	gtatcatttt	cttttttaac	cggtaagant	ttcagtttgt		300
agtaa	ctgtganaac	ccagtttccc	gtccatctcc	cttagggact	acccatagaa		360
aaagg	tccccacnga	agcaagaaga	taagtctttc	atggctgctg	gttgcttaaa		420
ctaaa	accaaaaaat	tccccttgga	aa				452

```
<210> 108
<211> 502
<212> DNA
<213> Homo sapien
```

```
<220>
<221> misc_feature
<222> (1)...(502)
<223> n = A,T,C or G
```

<210>	109
<211>	1308
<212>	DNA

<213> Homo sapien

<400> 109

```

accgaggtc tcgctaaaat catcatggat tcacttggcg ccgtcagcac tcgacttggg      60
tttgatcttt tcaaagagct gaagaaaaca aatgatggca acatcttctt ttccctctgt      120
ggcatcttga ctgcaattgg catggtcctc ctggggaccc gaggagccac cgcttcccag      180
ttggaggagg tgtttcactc tgaaaaagag acgaagagct caagaataaa ggctgaagaa      240
aaagaggtga ttgagaacac agaagcagta catcaacaat tccaaaagtt tttgactgaa      300
ataagcaaac tctaataatga ttatgaactg aacataacca acaggctgtt tggagaaaaa      360
acatacctct tccttcaaaa atacttagat tatgttgaaa aatattatca tgcctctctg      420
gaacctgttg attttgtaaa tgcagccgat gaaagtcgaa agaagattaa ttcctgggtt      480
gaaagcaaaa caaatgaaaa aatcaaggac ttgttcccag atggctctat tagtagctct      540
accaagctgg tgctggtgaa catggtttat tttaaagggc aatgggacag ggagtttaag      600
aaagaaaata ctaaggaaga gaaatttttg atgaataaga gcacaagtaa atctgtacag      660
atgatgacac agagccattc ctttagcttc actttcctgg aggacttgca ggccaaaatt      720
ctagggattc catataaaaa caacgacctc agcatgtttg tgcttctgcc caacgacatc      780
gatggcctgg agaagataat agataaaaata agtcctgaga aattggtaga gtggactagt      840
ccagggcata tggaagaaag aaaggtgaat ctgcacttgc cccgggttga ggtggaggac      900
agttacgata tagaggcggg cctggctgcc atggggatgg gcgatgcctt cagtgagcac      960
aaagccgact actcgggaat gtcgtcaggg tccgggttgt acgcccagaa gttcctgcac      1020
agttcctttg tggcagtaac tgaggaaggc accgaggctg cagctgccac tggcataggc      1080
tttactgtca catccgcccc aggtcatgaa aatgttcaact gcaatcatcc cttcctgttc      1140
ttcatcaggc acaatgaatc caacagcatc ctcttcttcg gcagattttc ttctccttaa      1200
gatgatcggt gccatggcat tgctgctttt agcaaaaaac aactaccagt gttactcata      1260
tgattatgaa aatcgtccat tcttttaaat ggtggctcac ttgcattt      1308

```

<210> 110

<211> 391

<212> PRT

<213> Homo sapien

<400> 110

```

Met Asp Ser Leu Gly Ala Val Ser Thr Arg Leu Gly Phe Asp Leu Phe
 1          5          10          15
Lys Glu Leu Lys Lys Thr Asn Asp Gly Asn Ile Phe Phe Ser Pro Val
 20          25          30
Gly Ile Leu Thr Ala Ile Gly Met Val Leu Leu Gly Thr Arg Gly Ala
 35          40          45
Thr Ala Ser Gln Leu Glu Glu Val Phe His Ser Glu Lys Glu Thr Lys
 50          55          60
Ser Ser Arg Ile Lys Ala Glu Glu Lys Glu Val Ile Glu Asn Thr Glu
 65          70          75          80
Ala Val His Gln Gln Phe Gln Lys Phe Leu Thr Glu Ile Ser Lys Leu
 85          90          95
Thr Asn Asp Tyr Glu Leu Asn Ile Thr Asn Arg Leu Phe Gly Glu Lys
 100         105         110
Thr Tyr Leu Phe Leu Gln Lys Tyr Leu Asp Tyr Val Glu Lys Tyr Tyr
 115         120         125
His Ala Ser Leu Glu Pro Val Asp Phe Val Asn Ala Ala Asp Glu Ser
 130         135         140
Arg Lys Lys Ile Asn Ser Trp Val Glu Ser Lys Thr Asn Glu Lys Ile
 145         150         155         160
Lys Asp Leu Phe Pro Asp Gly Ser Ile Ser Ser Ser Thr Lys Leu Val
 165         170         175

```

Leu Val Asn Met Val Tyr Phe Lys Gly Gln Trp Asp Arg Glu Phe Lys
 180 185 190
 Lys Glu Asn Thr Lys Glu Glu Lys Phe Trp Met Asn Lys Ser Thr Ser
 195 200 205
 Lys Ser Val Gln Met Met Thr Gln Ser His Ser Phe Ser Phe Thr Phe
 210 215 220
 Leu Glu Asp Leu Gln Ala Lys Ile Leu Gly Ile Pro Tyr Lys Asn Asn
 225 230 235 240
 Asp Leu Ser Met Phe Val Leu Leu Pro Asn Asp Ile Asp Gly Leu Glu
 245 250 255
 Lys Ile Ile Asp Lys Ile Ser Pro Glu Lys Leu Val Glu Trp Thr Ser
 260 265 270
 Pro Gly His Met Glu Glu Arg Lys Val Asn Leu His Leu Pro Arg Phe
 275 280 285
 Glu Val Glu Asp Ser Tyr Asp Leu Glu Ala Val Leu Ala Ala Met Gly
 290 295 300
 Met Gly Asp Ala Phe Ser Glu His Lys Ala Asp Tyr Ser Gly Met Ser
 305 310 315 320
 Ser Gly Ser Gly Leu Tyr Ala Gln Lys Phe Leu His Ser Ser Phe Val
 325 330 335
 Ala Val Thr Glu Glu Gly Thr Glu Ala Ala Ala Thr Gly Ile Gly
 340 345 350
 Phe Thr Val Thr Ser Ala Pro Gly His Glu Asn Val His Cys Asn His
 355 360 365
 Pro Phe Leu Phe Phe Ile Arg His Asn Glu Ser Asn Ser Ile Leu Phe
 370 375 380
 Phe Gly Arg Phe Ser Ser Pro
 385 390

<210> 111
 <211> 1419
 <212> DNA
 <213> Homo sapien

<400> 111

ggagaactat	aaattaagga	tcccagctac	ttaattgact	tatgcttcct	agttcgttgc	60
ccagccacca	ccgtctctcc	aaaaaccga	ggtctcgcta	aaatcatcat	ggattcactt	120
ggcgccgtca	gcactcgact	tgggtttgat	cttttcaaag	agctgaagaa	aacaaatgat	180
ggcaacatct	tcttttcccc	tgtgggcata	ttgactgcaa	ttggcatggt	cctcctgggg	240
acccgaggag	ccaccgcttc	ccagttggag	gaggtgtttc	actctgaaaa	agagacgaag	300
agctcaagaa	taaaggctga	agaaaaagag	gtggttaagaa	taaaggctga	aggaaaagag	360
attgagaaca	cagaagcagt	acatcaacaa	ttccaaaagt	ttttgactga	aataagcaaa	420
ctcactaatg	attatgaact	gaacataacc	aacaggctgt	ttggagaaaa	aacatacctc	480
ttccttcaaa	aatacttaga	ttatgttgaa	aaatattatc	atgcatctct	ggaacctggt	540
gattttgtaa	atgcagccga	tgaaagtcga	aagaagatta	attcctgggt	tgaaagcaaa	600
acaaatgaaa	aatcaagga	cttggtccca	gatggctcta	ttagtagctc	taccaagctg	660
gtgctggtga	acatggttta	ttttaaaggg	caatgggaca	gggagtttaa	gaaagaaaat	720
actaaggaag	agaaattttg	gatgaataag	agcacaagta	aatctgtaca	gatgatgaca	780
cagagccatt	ccttttagctt	cactttcctg	gaggacttgc	aggccaaaat	tctagggatt	840
ccatataaaa	acaacgacct	aagcatgttt	gtgcttctgc	ccaacgacat	cgatggcctg	900
gagaagataa	tagataaaat	aagtctctgag	aaattggtag	agtggactag	tccagggcat	960
atggaagaaa	gaaagggtgaa	tctgcacttg	ccccggtttg	aggtggagga	cagttacgat	1020
ctagaggcg	tcctggctgc	catgggggatg	ggcagtgctt	tcagtgaagca	caaagccgac	1080
tactcgggaa	tgctgtcagg	ctccgggttg	tacgccacaga	agttcctgca	cagttccttt	1140


```
<210> 112
<211> 400
<212> PRT
<213> Homo sapien
```

Met 1	Asp	Ser	Leu	Gly 5	Ala	Val	Ser	Thr	Arg 10	Leu	Gly	Phe	Asp	Leu 15	Phe
Lys	Glu	Leu	Lys 20	Lys	Thr	Asn	Asp	Gly 25	Asn	Ile	Phe	Phe	Ser 30	Pro	Val
Gly	Ile	Leu 35	Thr	Ala	Ile	Gly	Met 40	Val	Leu	Leu	Gly	Thr 45	Arg	Gly	Ala
Thr	Ala 50	Ser	Gln	Leu	Glu	Glu 55	Val	Phe	His	Ser	Glu 60	Lys	Glu	Thr	Lys
Ser 65	Ser	Arg	Ile	Lys	Ala 70	Glu	Glu	Lys	Glu	Val 75	Val	Arg	Ile	Lys	Ala 80
Glu	Gly	Lys	Glu 85	Ile	Glu	Asn	Thr	Glu	Ala 90	Val	His	Gln	Gln 95	Phe	Gln
Lys	Phe	Leu	Thr 100	Glu	Ile	Ser	Lys	Leu 105	Thr	Asn	Asp	Tyr	Glu 110	Leu	Asn
Ile	Thr	Asn 115	Arg	Leu	Phe	Gly	Glu 120	Lys	Thr	Tyr	Leu	Phe 125	Leu	Gln	Lys
Tyr	Leu 130	Asp	Tyr	Val	Glu	Lys 135	Tyr	Tyr	His	Ala	Ser 140	Leu	Glu	Pro	Val
Asp 145	Phe	Val	Asn	Ala	Ala 150	Asp	Glu	Ser	Arg	Lys 155	Lys	Ile	Asn	Ser	Trp 160
Val	Glu	Ser	Lys 165	Thr	Asn	Glu	Lys	Ile	Lys 170	Asp	Leu	Phe	Pro	Asp 175	Gly
Ser	Ile	Ser	Ser 180	Ser	Thr	Lys	Leu	Val 185	Leu	Val	Asn	Met	Val 190	Tyr	Phe
Lys	Gly	Gln 195	Trp	Asp	Arg	Glu	Phe 200	Lys	Lys	Glu	Asn 205	Thr	Lys	Glu	Glu
Lys	Phe 210	Trp	Met	Asn	Lys	Ser 215	Thr	Ser	Lys	Ser	Val 220	Gln	Met	Met	Thr
Gln 225	Ser	His	Ser	Phe	Ser 230	Phe	Thr	Phe	Leu	Glu 235	Asp	Leu	Gln	Ala	Lys 240
Ile	Leu	Gly	Ile 245	Pro	Tyr	Lys	Asn	Asn	Asp 250	Leu	Ser	Met	Phe	Val 255	Leu
Leu	Pro	Asn	Asp 260	Ile	Asp	Gly	Leu	Glu 265	Lys	Ile	Ile	Asp	Lys 270	Ile	Ser
Pro	Glu	Lys 275	Leu	Val	Glu	Trp	Thr 280	Ser	Pro	Gly	His	Met	Glu 285	Glu	Arg
Lys	Val 290	Asn	Leu	His	Leu	Pro 295	Arg	Phe	Glu	Val	Glu 300	Asp	Ser	Tyr	Asp
Leu 305	Glu	Ala	Val	Leu	Ala 310	Ala	Met	Gly	Met	Gly 315	Asp	Ala	Phe	Ser	Glu 320
His	Lys	Ala	Asp 325	Tyr	Ser	Gly	Met	Ser	Ser 330	Gly	Ser	Gly	Leu	Tyr 335	Ala

```

Gln Lys Phe Leu His Ser Ser Phe Val Ala Val Thr Glu Glu Gly Thr
      340                      345                      350
Glu Ala Ala Ala Ala Thr Gly Ile Gly Phe Thr Val Thr Ser Ala Pro
      355                      360                      365
Gly His Glu Asn Val His Cys Asn His Pro Phe Leu Phe Phe Ile Arg
      370                      375                      380
His Asn Glu Ser Asn Ser Ile Leu Phe Phe Gly Arg Phe Ser Ser Pro
      385                      390                      395                      400

```

```

<210> 113
<211> 957
<212> DNA
<213> Homo sapien

```

```
<400> 113
```

```

ctcgaccttc tctgcacagc ggatgaaccc tgagcagctg aagaccagaa aagccactat      60
gactttctgc ttaattcagg agcttacagg attcttcaaa gagtgtgtcc agcatccttt      120
gaaacatgag ttcttaccag cagaagcaga cctttacccc accacctcag cttcaacagc      180
agcaggtgaa acaacccagc cagcctccac ctcaggaaat atttgttccc acaaccaagg      240
agccatgcc acaaaagggt ccacaacctg gaaacacaaa gattccagag ccaggctgta      300
ccaaggtccc tgagccaggc tgtaccaagg tccctgagcc aggttgtacc aagggtccctg      360
agccaggatg taccaaggtc cctgagccag gttgtacca ggtccctgag ccaggctaca      420
ccaaggtccc tgagccaggc agcatcaagg tccctgacca aggtctcatc aagtttcctg      480
agccagggtgc catcaaagtt cctgagcaag gatacaccaa agttcctgtg ccaggctaca      540
caaaggtacc agagccatgt ccttcaacgg tcaactccagg cccagctcag cagaagacca      600
agcagaagta atttggtgca cagacaagcc cttgagaagc caaccaccag atgctggaca      660
ccctcttccc atctgtttct gtgtcttaat tgtctgtaga ccttgtaatc agtacattct      720
caccccaagc catagtctct ctcttatttg tctcctaaaa atacggtact ataaagcttt      780
tgttcacaca cactctgaag aatcctgtaa gccctgaat taagcagaaa gtcttcatgg      840
cttttctggt cttcggtctgc tcagggttca tctgaagatt cgaatgaaaa gaaatgcatg      900
tttctgtctc tgccctcatt aaattgcttt taattccaaa aaaaaaaaaa aaaaaaaa      957

```

```

<210> 114
<211> 161
<212> PRT
<213> Homo sapien

```

```
<400> 114
```

```

Met Ser Ser Tyr Gln Gln Lys Gln Thr Phe Thr Pro Pro Pro Gln Leu
  1          5          10          15
Gln Gln Gln Gln Val Lys Gln Pro Ser Gln Pro Pro Pro Gln Glu Ile
      20          25          30
Phe Val Pro Thr Thr Lys Glu Pro Cys His Ser Lys Val Pro Gln Pro
      35          40          45
Gly Asn Thr Lys Ile Pro Glu Pro Gly Cys Thr Lys Val Pro Glu Pro
      50          55          60
Gly Cys Thr Lys Val Pro Glu Pro Gly Cys Thr Lys Val Pro Glu Pro
      65          70          75          80
Gly Cys Thr Lys Val Pro Glu Pro Gly Cys Thr Lys Val Pro Glu Pro
      85          90          95
Gly Tyr Thr Lys Val Pro Glu Pro Gly Ser Ile Lys Val Pro Asp Gln
      100          105          110
Gly Phe Ile Lys Phe Pro Glu Pro Gly Ala Ile Lys Val Pro Glu Gln
      115          120          125

```

Gly Tyr Thr Lys Val Pro Val Pro Gly Tyr Thr Lys Val Pro Glu Pro
 130 135 140
 Cys Pro Ser Thr Val Thr Pro Gly Pro Ala Gln Gln Lys Thr Lys Gln
 145 150 155 160
 Lys

<210> 115
 <211> 506
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(506)
 <223> n = A,T,C or G

<400> 115

cattggtnc	ttcatttgc	ntggaagtgt	nnatctctaa	cagtggacaa	agttcccngt	60
gccttaaa	ctgtnacac	tttggaant	gaaaantng	tantatgata	ggttattctg	120
angtanag	gttctggata	ccattanatn	tgccccngt	gtcagaggct	catattgtgt	180
tatgtaaa	gtatntcatt	cgctactatn	antcaattng	aaatanggtc	tttgggttat	240
gaatantnng	cagcncanct	nanangctgt	ctgtngtatt	cattgtggtc	atagcacctc	300
acancattgt	aacctcnatc	nagttagaca	nactagna	aan	ttcctagtga	360
ttccaaatgg	ntcatntcn	aatgtttaa	agttanttaa	gtgtaagaaa	tacagactgg	420
atgttccacc	aactagtacc	tgtaatgacn	ggcctgtccc	aacacatctc	ccttttccat	480
gactgtggta	nccgcgcatg	gaaaaa				506

<210> 116
 <211> 3079
 <212> DNA
 <213> Homo sapien

<400> 116

ggatccccg	gtttcctaaa	ccccccacag	agtccctgcc	aggccaaaga	gcaaggaaaa	60
ggtcaaaagg	cagaaaaaat	gctgagttag	gaggagctat	ggaaggataa	acctggcctt	120
aaagagggtca	aagtggttta	tagggggcgc	tgagggtctc	ccacattctc	tggcctaaac	180
cttgacaggca	gatctgccca	gtgggctctg	ggatagctgt	gccttcccta	acaaaaaat	240
tgtgcacaaa	aggatgaaac	tctattttcc	ctctagcaca	taaccaagaa	tataaggcta	300
cagattgcct	ttcccagagg	gaaaaccctg	cagcaaccctg	ctgcctggaa	aagtgtgaaga	360
gcagatcact	ggggaatcgt	ttgccccccg	ctgatggaca	gcttccccaa	gctccaaggg	420
cagggtgctca	gcatgtaccg	tactgggatg	gttgtaata	ctcctggtcc	tgtaagagtc	480
ccaggacact	gccatgcca	tgccccctca	gttccctggca	tccttttttg	gctgctcaca	540
gccccagcct	ctatggtgaa	gacatacttg	ctagcagcgt	caccaacttg	ttgccaaagag	600
atcagtgtc	gaaggcaagg	ttattttctaa	ctgagcagag	cctgccagga	agaaagcggt	660
tgcaccccac	accactgtgc	aggtgtgacc	ggtgagctca	cagctgcccc	ccaggcatgc	720
ccagcccact	taatcatcac	agctcgacag	ctctctcgcc	cagcccagtt	ctggaaggga	780
taaaaagggg	catcaccggt	cctgggtaac	agagccacct	tctgctcct	gctgagctct	840
gttctctcca	gcacctccca	accactagt	gacctgttct	cttgctccac	caggaacaag	900
ccaccatgtc	tcgccagtca	agtgtgtctt	ccgggagcggg	gggcagtcgt	agcttcagca	960
ccgctctgc	catcaccccg	tctgtctccc	gcaccagctt	cacctccgtg	tcccggctccg	1020
gggggtggcg	tgggtgggtgc	ttcggcaggg	tcagccttgc	gggtgcttgt	ggagtgggtg	1080
gctatggcag	ccggagcctc	tacaacctgg	ggggctccaa	gaggatatcc	atcagcacta	1140
gtgggtggcag	cttcaggaac	cggtttggtg	ctgggtgctgg	aggcggctat	ggctttggag	1200

gtggtgccgg	tagtggattt	ggtttcggcg	gtggagctgg	tggtggcttt	gggctcggtg	1260
gcgagactgg	ctttggaggt	ggcttcgggtg	gccctggctt	tcctgtctgc	cctcctggag	1320
gtatccaaga	ggctactgtc	aaccagagtc	tcctgactcc	cctcaacctg	caaatcgacc	1380
ccagcatcca	gaggggtgagg	accgaggagc	gcgagcagat	caagaccctc	aacaataagt	1440
ttgctctctt	catcgacaag	gtgcggttcc	tggagcagca	gaacaagggt	ctggaaacaa	1500
agtggaccct	gctgcaggag	cagggcacca	agactgtgag	gcagaacctg	gagccgttgt	1560
tcgagcagta	catcaacaac	ctcaggaggc	agctggacag	catcgtgggg	gaacggggcc	1620
gcctggactc	agagctgaga	aacatgcagg	acctggtgga	agacttcaag	aacaagtatg	1680
aggatgaaat	caacaagcgt	accactgctg	agaatgagtt	tgtgatgctg	aagaaggatg	1740
tagatgctgc	ctacatgaac	aaggtggagc	tggaggccaa	ggttgatgca	ctgatggatg	1800
agattaactt	catgaagatg	ttctttgatg	cggagctgtc	ccagatgcag	acgcatgtct	1860
ctgacacctc	agtggctctc	tccatggaca	acaaccgcaa	cctggacctg	gatagcatca	1920
tcgctgaggt	caaggcccag	tatgaggaga	ttgccaaccg	cagccggaca	gaagccgagt	1980
cctggtatca	gaccaagtat	gaggagctgc	agcagacagc	tggccggcat	ggcgatgacc	2040
tccgcaacac	caagcatgag	atctctgaga	tgaaccggat	gatccagagg	ctgagagccg	2100
agattgacaa	tgtcaagaaa	cagtgcgcca	atctgcagaa	cgccattgctg	gatgccgagc	2160
agcgtgggga	gctggccctc	aaggatgcca	ggaacaagct	ggccgagctg	gaggaggccc	2220
tgcagaaggc	caagcaggac	atggcccggc	tgtgcgtga	gtaccaggag	ctcatgaaca	2280
ccaagctggc	cctggacgtg	gagatcgcca	cttaccgcaa	gctgctggag	ggcgagggaat	2340
gcagactcag	tggagaagga	gttggaccag	tcaacatctc	tgttgtcaca	agcagtgttt	2400
cctctggata	tggcagtggc	agtggctatg	gcggtggcct	cgggtggaggt	cttggcggcg	2460
gcctcgggtg	aggtcttggc	ggaggtagca	gtggaagcta	ctactccagc	agcagtgggg	2520
gtgtcggcct	aggtggtggg	ctcagtgtgg	ggggctctgg	cttcagtgc	agcagtagcc	2580
gagggctggg	ggtgggcttt	ggcagtggcg	ggggtagcag	ctccagcgtc	aaatttgtct	2640
ccaccacctc	ctctcccg	aagagcttca	agagctaaga	acctgctgca	agtcaactgcc	2700
ttccaagtgc	agcaaccacg	cccatggaga	ttgctctctc	taggcagttg	ctcaagccat	2760
gttttatcct	tttctggaga	gtagtctaga	ccaagccaat	tgcagaacca	cattctttgg	2820
ttcccaggag	agccccattc	ccagcccctg	gtctcccgtg	ccgcagttct	atattctgct	2880
tcaaatcagc	cttcagggtt	cccacagcat	ggcccctgct	gacacgagaa	cccaaagttt	2940
tcccaaatct	aaatcatcaa	aacagaatcc	ccaccccaat	cccaaatttt	gttttggttc	3000
taactacctc	cagaatgtgt	tcaataaaat	gtttttataat	ataagctggg	gtgcagaatt	3060
gttttttttt	tetacecaa					3079

<210> 117

<211> 6921

<212> DNA

<213> Homo sapien

<400> 117

gaattctgac	tgtccactca	aaactttctat	tccgatcaaa	gctatctgtg	actacagaca	60
aattgagata	accattttaca	aagacgatga	atgtgttttg	gcgaataact	ctcatcgtgc	120
taaatggaag	gtcattagtc	ctactgggaa	tgaggctatg	gtcccatctg	tgtgcttcac	180
cgttcctcca	ccaaacaaag	aagcggtgga	ccttgccaac	agaattgagc	aacagtatca	240
gaatgtcctg	actctttggc	atgagtctca	cataaacatg	aagagtgtag	tatcctggca	300
ttatctcatc	aatgaaattg	atagaattcg	agctagcaat	gtggcttcaa	taaagacaat	360
gtacacctgg	gaacatcagc	aagttctaag	taatctacaa	tctcgttttg	aagattttct	420
ggaagatagc	caggaatccc	aagtcttttc	aggctcagat	ataacacaa	tggaaaagga	480
ggttaatgta	tgtaaagcag	attatcaaga	acttcttaaa	tctgcagaaa	gagaggagca	540
agaggaatca	gtttataatc	tctacatctc	tgaagttcga	aacattagac	ttcggttaga	600
gaactgtgaa	gacgggctga	ttagacagat	tcgaactccc	ctggaaagag	atgatttgca	660
tgaaagtgtg	ttcagaatca	cagaacagga	gaaactaaag	aaagagctgg	aacgacttaa	720
agatgatttg	ggaacaatca	caaataagtg	tgaggagttt	ttcagtcaag	cagcagcctc	780
ttcatcagtc	cctaccctac	gatcagagct	taatgtggtc	cttcagaaca	tgaaccaagt	840
ctattctatg	tcttccactt	acatagataa	gttgaaaact	gttaacttgg	tgttaaaaaa	900

cactcaagct	gcagaagccc	tcgtaaaact	ctatgaaact	aaactgtgtg	aagaagaagc	960
agttatagct	gacaagaata	atattgagaa	tctaataagt	actttaagc	aatggagatc	1020
tgaagtagat	gaaaagagac	aggtattcca	tgcccttagag	gatgagttgc	agaaagctaa	1080
agccatcagt	gatgaaatgt	ttaaaacgta	taaagaacgg	gaccttgatt	ttgactggca	1140
caaagaaaaa	gcagatcaat	tagttgaaag	gtggcaaaat	gttcatgtgc	agattgacaa	1200
caggttacgg	gacttagagg	gcattggcaa	atcactgaag	tactacagag	acacttacca	1260
tccttttagat	gattggatcc	agcaggttga	aactactcag	agaaagattc	aggaaaatca	1320
gcctgaaaat	agtaaaaccc	tagccacaca	gttgaatcaa	cagaagatgc	tggtgtccga	1380
aatagaaatg	aaacagagca	aaatggacga	gtgtcaaaaa	tatgcagaac	agtactcagc	1440
tacagtgaag	gactatgaat	tacaaacaat	gacctaccgg	gccatggtag	attcacaaca	1500
aaaatctcca	gtgaaacgcc	gaagaatgca	gagttcagca	gatctcatta	ttcaagagtt	1560
catggaccta	aggactcgat	atactgccct	ggtcactctc	atgacacaat	atatttaaatt	1620
tgctggtgat	tcattgaaga	ggctggaaga	ggaggagatt	aaaaggtgta	aggagacttc	1680
tgaacatggg	gcataattcag	atctgcttca	gcgtcagaag	gcaacagtgc	ttgagaatag	1740
caaacttaca	ggaaagataa	gtgagttgga	aagaatggta	gctgaactaa	agaaacaaaa	1800
gtcccagagta	gaggaagaac	ttccgaaggt	cagggaggct	gcagaaaatg	aattgagaaa	1860
gcagcagaga	aatgtagaag	atatctctct	gcagaagata	agggctgaaa	gtgaagccaa	1920
gcagtaccgc	agggaaacttg	aaaccattgt	gagagagaag	gaagccgctg	aaagagaact	1980
ggagcgggtg	aggcagctca	ccatagaggg	cgaggctaaa	agagctgccg	tggaagagaa	2040
cctcctgaat	tttcgcaatc	agttggagga	aaacaccttt	accagacgaa	cactggaaga	2100
tcctcttaaa	agaaaagatt	taagtctcaa	tgattttggag	caacaaaaaa	ataaattaat	2160
ggaagaatta	agaagaaaga	gagacaatga	ggaagaactc	ttgaagctga	taaagcagat	2220
ggaaaaagac	cttgcatthc	agaaacaggt	agcagagaaa	cagttgaaag	aaaagcagaa	2280
aattgaattg	gaagcaagaa	gaaaaataac	tgaaattcag	tatacatgta	gagaaaatgc	2340
attgccagtg	tgtccgatca	cacaggctac	atcatgcagg	gcagtaacgg	gtctccagca	2400
agaacatgac	aagcagaaag	cagaagaact	caaacagcag	gtagatgaac	taacagctgc	2460
caatagaaag	gctgaacaag	acatgagaga	gctgacatat	gaacttaatg	ccctccagct	2520
tgaaaaaacg	tcattctgagg	aaaaggctcg	tttgctaaaa	gataaactag	atgaaacaaa	2580
taatacactc	agatgcctta	agttggagct	ggaaagggaag	gatcaggcgg	agaaagggtg	2640
ttctcaacaa	ctcagagagc	ttggtaggca	attgaatcaa	accacaggta	aagctgaaga	2700
agccatgcaa	gaagctagtg	atctcaagaa	aataaagcgc	aattatcagt	tagaattaga	2760
atctcttaat	catgaaaaag	ggaaactaca	aagagaagta	gacagaatca	caagggcaca	2820
tgctgtagct	gagaagaata	ttcagcattt	aaattcacaa	attcattctt	ttcgagatga	2880
gaaagaatta	gaaagactac	aaatctgcca	gagaaaatca	gatcatctaa	aagaacaatt	2940
tgagaaaaagc	catgagcagt	tgcttcaaaa	tatcaaagct	gaaaaagaaa	ataatgataa	3000
aatccaaagg	ctcaatgaag	aattggagaa	aagtaatgag	tgtgcagaga	tgctaaaaca	3060
aaaagtagag	gagcttacta	ggcagaataa	tgaaccacaa	ttaatgatgc	agagaattca	3120
ggcagaatca	gagaatatag	tttttagagaa	acaaactatc	cagcaaagat	gtgaagcact	3180
gaaaattcag	gcagatgggt	ttaaagatca	gctacgcagc	acaaatgaac	acttgcataa	3240
acagacaaaa	acagagcagg	atthttcaaag	aaaaatttaa	tgcttagaag	aagacctggc	3300
gaaaagtcaa	aatttggttaa	gtgaatttaa	gcaaaagtgt	gaccaacaga	acattatcat	3360
ccagaatacc	aagaaagaag	ttagaaatct	gaatgcggaa	ctgaatgctt	ccaaagaaga	3420
gaagcgacgc	ggggagcaga	aagttcagct	acaacaagct	caggtgcaag	agttaaataa	3480
caggttgaaa	aaagtacaag	acgaattaca	cttaaagacc	atagaggagc	agatgacca	3540
cagaaagatg	gttctgtttc	aggaagaatc	tggtaaattc	aaacaatcag	cagaggagtt	3600
tcggaagaag	atggaaaaat	taatggagtc	caaagtcac	actgaaaatg	atatttcagg	3660
cattaggtct	gacttttgtgt	ctcttcaaca	agaaaactct	agagcccaag	aaaatgctaa	3720
gctttgtgaa	acaaacatta	aagaacttga	aagacagctt	caacagtatc	gtgaacaaat	3780
gcagcaaggg	cagcacatgg	aagcaaatca	ttaccaaaaa	tgtcagaaac	ttgaggatga	3840
gctgatagcc	cagaagcgtg	aggttgaaaa	cctgaagcaa	aaaatggacc	aacagatcaa	3900
agagcatgaa	catcaattag	ttttgctcca	gtgtgaaatt	caaaaaaaga	gcacagccaa	3960
agactgtacc	ttcaaacaccg	attttgagat	gacagtgaag	gagtgccagc	actctggaga	4020
gctgtcctct	agaaacactg	gacccttcca	cccaacacc	agatccctc	tggtgagatg	4080
gactcaagaa	ccacagccat	tggaagagaa	gtggcagcat	cgggttggtg	aacagatacc	4140

caaagaagtc	caattccagc	caccaggggc	tccactcgag	aaagagaaaa	gccagcagtg	4200
ttactctgag	tactttttctc	agacaagcac	cgagttacag	ataacttttg	atgagacaaa	4260
ccccattaca	agactgtctg	aaattgagaa	gataagagac	caagccctga	acaattctag	4320
accacctgtt	aggtatcaag	ataacgcatg	tgaatggaa	ctggtgaagg	ttttgacacc	4380
cttagagata	gctaagaaca	agcagtatga	tatgcataca	gaagtcacaa	cattaaaaaca	4440
agaaaagaac	ccagttccca	gtgctgaaga	atggatgctt	gaagggtgca	gagcatctgg	4500
tggactcaag	aaaggggatt	tccttaagaa	gggcttagaa	ccagagacct	tccagaactt	4560
tgatggatgat	catgcatggt	cagtcaggga	tgatgaattt	aaattccaag	ggcttaggca	4620
cactgtgact	gccaggcagt	tgggtggaagc	taagcttctg	gacatgagaa	caattgagca	4680
gctgcgactc	ggtcttaaga	ctggtgaaga	agttcagaaa	actcttaaca	agtttctgac	4740
gaaagccacc	tcaattgcag	ggctttacct	agaatctaca	aaagaaaaga	tttcatttgc	4800
ctcagcggcc	gagagaatca	taatagacaa	aatggtggct	ttggcatttt	tagaagctca	4860
ggctgcaaca	ggttttataa	ttgatcccat	ttcaggtcag	acatattctg	ttgaagatgc	4920
agttcttaaa	ggagtgtgtg	accccgcaatt	cagaattagg	cttcttgagg	cagagaaggc	4980
agctgtggga	tattcttatt	cttctaagac	attgtcagtg	tttcaagcta	tggaaaatag	5040
aatgcttgac	agacaaaaag	gtaaacatat	cttgggaagc	cagattgcca	gtgggggtgt	5100
cattgaccct	gtgagaggca	ttcgtgttcc	tccagaaatt	gctctgcagc	aggggttgtt	5160
gaataatgcc	atcttacagt	ttttacatga	gccatccagc	aacacaagag	ttttccctaa	5220
tccaataaac	aagcaagctc	tgtattactc	agaattactg	cgaatgtgtg	tatttgatgt	5280
agagtcccaa	tgtcttctgt	ttccatttgg	ggagaggaa	atttccaatc	tcaatgtcaa	5340
gaaaacacat	agaatttctg	tagtagatac	taaaacagga	tcagaattga	cogtgtatga	5400
ggctttccag	agaaacctga	ttgagaaaag	tatatatctt	gaactttcag	ggcagcaata	5460
tcagtgggaag	gaagctatgt	tttttgaatc	ctatgggcat	tcttctcata	tgtgtactga	5520
tactaaaaca	ggattacact	tcaatattaa	tgaggctata	gagcagggaa	caattgacaa	5580
agccttggtc	aaaaagtatc	aggaaggcct	catcacactt	acagaacttg	ctgattcttt	5640
gctgagccgg	ttagtcccca	agaaagattt	gcacagtcct	gttgccagggt	attggctgac	5700
tgctagtggg	gaaaggatct	ctgtactaaa	agcctcccgt	agaaatttgg	ttgatcggat	5760
tactgccttc	cgatgccttg	aagcccaagt	cagtacaggg	ggcataattg	atcctcttac	5820
tggcaaaaag	taccgggtgg	ccgaagcttt	gcatagaggc	ctggttgatg	aggggtttgc	5880
ccagcagctg	cgacagtgtg	aattagtaat	cacagggtt	ggccatccca	tcactaacaa	5940
aatgatgtca	gtggtggaag	ctgtgaatgc	aaatattata	aataaggaaa	tgggaatccg	6000
atgtttggaa	tttcagtact	tgacaggagg	gttgatagag	ccacagggtt	actctcgggt	6060
atcaatagaa	gaggctctcc	aagtaggtat	tatagatgtc	ctcattgcca	caaaaactcaa	6120
agatcaaaaag	tcatatgtca	gaaatataat	atgccctcag	acaaaaagaa	agttgacata	6180
taaagaagcc	ttagaaaaag	ctgatttttga	tttccacaca	ggacttaaac	tgttagaagt	6240
atctgagccc	ctgatgacag	gaatttctag	cctctactat	tcttccta	gggacatggt	6300
taaataactg	tgcaaggggt	gatgcaggct	ggttcatgcc	actttttcag	agtatgatga	6360
tatcggtctac	atatgcagtc	tgtgaattat	gtaacatact	ctatttcttg	agggctgcaa	6420
attgctaagt	gtcaaaaata	gagtaagttt	taaattgaaa	attacataag	atttaatgcc	6480
cttcaaatgg	tttcattttag	ccttgagaat	ggttttttga	aacttggcca	cactaaaatg	6540
tttttttttt	tttacgtaga	atgtgggata	aacttgatga	actccaagtt	cacagtgtca	6600
tttcttcaga	actcccttcc	attgaatagt	gatcatttat	taaatgataa	attgcactcg	6660
ctgaaagagc	acgtcatgaa	gcaccatgga	atcaaagaga	aagatataaa	ttcgttccca	6720
cagccttcaa	gctgcagtgt	tttagattgc	ttcaaaaaat	gaaaaagttt	tgcccttttcc	6780
gatatatagtga	ccttcttttgc	atattaaaaat	gtttaccaca	atgtcccatt	tctagttaag	6840
tcttcgcact	tgaagactaa	cattatgaat	attatgtgtt	ggaggagggg	aaggattttc	6900
ttcattctgt	gtatttttccg	g				6921

<210> 118

<211> 946

<212> DNA

<213> Homo sapien

<400> 118

```
<210> 119
<211> 8948
<212> DNA
<213> Homo sapien
```

tcaacagccc	ctgctccttg	ggccctcca	tgccatgccg	taatctctcc	cacccgacca	60
acaccaacac	ccagctccga	cgcagctcct	ctgcccctt	gccgccctcc	gagccacagc	120
tttctcccc	ctcctgcccc	cgcccgtcg	ccgtctccgc	gctcgacagc	gcctcgggag	180
ggcccaggtg	gcgagcagcg	acctcgcgag	ccttcgcac	tcccgcccg	ttccccggcc	240
gtccgcctat	ccttgccccc	ctccgctttc	tccgcgcgg	ccgcctcgc	ttatgcctcg	300
gcgctgagcc	gctctccga	ttgcccgcg	acatgagctg	caacggaggc	tcccaccgc	360
ggatcaacac	tctgggccgc	atgatccgcg	ccgagctctg	ccggacctg	cgctacgagg	420
tgaccagcgg	cggcgggggc	accagcagga	tgtactattc	tggcgcggc	gtgatcaccg	480
accagaactc	ggacggctac	tgtcaaaccc	gcacgatgtc	caggcaccag	aaccagaaca	540
ccatccagga	gctgctgcag	aactgctccg	actgcttgat	gcgagcagag	ctcatcgtgc	600
agcctgaatt	gaagtatgga	gatggaatac	aactgactcg	gagtcgagaa	ttggatgagt	660
gttttgccca	ggccaatgac	caaatggaaa	tcttcgacag	cttgatcaga	gagatgcggc	720
agatgggcca	gccctgtgat	gcttaccaga	aaaggctctc	tcagctccaa	gagcaaatgc	780
gagcccttta	taaagccatc	agtgtccctc	gagtcgcgag	ggcagctccc	aagggtgggtg	840
gaggctacac	ttgtcagagt	ggctctggct	gggatgagtt	caccaaacat	gtccaggatg	900
aatgtttggg	gtggatgagg	cagcaaaagg	cggagatgga	catggtggcc	tggggtgtgg	960
acctggcctc	agtggagcag	cacattaaca	gccaccgggg	catccacaac	tccatcggcg	1020
actatcgctg	gcagctggac	aaaatcaaag	ccgacctcgc	cgagaaatct	gcgatctacc	1080
agttggagga	ggagtatgaa	aacctgctga	aagcgtcctt	tgagaggatg	gatcacctgc	1140
gacagctgca	gaacatcatt	caggccacgt	ccaggggagat	catgtggatc	aatgactgcg	1200
aggaggagga	gctgctgtac	gactggagcg	acaagaacac	caacatcgct	cagaaacagg	1260
aggcctttct	catacgcatg	agtcaactgg	aagttaaaga	aaaagagctc	aataagctga	1320
aacaagaaag	tgaccaactt	gtcctcaatc	agcatccagc	ttcagacaaa	attgaggcct	1380
atatggacac	tctgcagacg	cagtggagtt	ggattcttca	gatcaccaag	tgcattgatg	1440
tcatctgaa	agaaaatgct	gctacttttc	agttttttga	agaggcgcag	tctactgaag	1500
catacctgaa	ggggctccag	gactccatca	ggaagaagta	cccctgcgac	aagaacatgc	1560
ccctgcagca	cctgctgga	cagatcaagg	agctggagaa	agaacgagag	aaaatccttg	1620
aatacaagcg	tcagggtgcag	aacttggtaa	acaagtctaa	gaagattgta	cagctgaagc	1680
ctcgtaaccc	agactacaga	agcaataaac	ccattattct	cagagctctc	tgtgactaca	1740
aacaagatca	gaaaatcgtg	cataaggggg	atgagtgtat	cctgaaggac	aacaacgagc	1800
gcagcaagtg	gtacgtgacg	ggcccggggg	gcgttgacat	gcttggtccc	tctgtggggc	1860

tgatcatccc	tcctccgaac	ccactggccg	tggacctctc	ttgcaagatt	gagcagtact	1920
acgaagccat	cttggtctctg	tggaaccagc	tctacatcaa	catgaagagc	ctggtgtcct	1980
ggcactactg	catgattgac	atagagaaga	tcagggccat	gacaatcgcc	aagctgaaaa	2040
caatgcggca	ggaagattac	atgaagacga	tagccgacct	tgagttacat	taccaagagt	2100
tcatcagaaa	tagccaaggc	tcagagatgt	ttggagatga	tgacaagcgg	aaaatacagt	2160
ctcagttcac	cgatgcccag	aagcattacc	agaccctggt	cattcagctc	cctggctatc	2220
cccagcacca	gacagtgacc	acaactgaaa	tcactcatca	tggaaacctgc	caagatgtca	2280
accataataa	agtaattgaa	accaacagag	aaaatgacaa	gcaagaaaca	tggatgctga	2340
tggagctgca	gaagattcgc	aggcagatag	agcactgcga	gggcaggatg	actctcaaaa	2400
acctccctct	agcagaccag	gggtctttct	accacatcac	agtgaaaatt	aacgagctta	2460
agagtgtgca	gaatgattca	caagcaattg	ctgaggttct	caaccagctt	aaagatatgc	2520
ttgccaaact	cagaggttct	gaaaagtact	gctattttaca	gaatgaagta	tttggactat	2580
ttcagaaaact	ggaaaatatc	aatggtgtta	cagatggcta	cttaaatagc	ttatgcacag	2640
taagggcact	gctccaggct	attctccaaa	cagaagacat	gttaaagggt	tatgaagcca	2700
ggctcactga	ggaggaaact	gtctgcctgg	acctggataa	agtggaaagt	taccgctgtg	2760
gactgaagaa	aataaaaaat	gacttgaact	tgaagaagtc	gttggtggcc	actatgaaga	2820
cagaactaca	gaaagcccag	cagatccact	ctcagacttc	acagcagtat	ccactttatg	2880
atctggactt	gggcaagttc	ggtgaaaaag	tcacacagct	gacagaccgc	tggcaaagga	2940
tagataaaca	gatcgacttt	agattatggg	acctggagaa	acaaatcaag	caattgagga	3000
attatcgtga	taactatcag	gctttctgga	agtgcgtcta	tgatcgtaaa	cgccgcagg	3060
attccttaga	atccatgaaa	tttggagatt	ccaacacagt	catgcggttt	ttgaatgagc	3120
agaagaactt	gcacagtga	atatctggca	aacgagacaa	atcagaggaa	gtacaaaaaa	3180
ttgctgaact	ttgcgccaat	tcaattaagg	attatgagct	ccagctggcc	tcatacacct	3240
caggactgga	aactctgctg	aacataccta	tcaagaggac	catgattcag	tccccttctg	3300
gggtgattct	gcaagaggct	gcagatgttc	atgctcggta	cattgaaacta	cttacaagat	3360
ctggagacta	ttacagggtc	ttaagtgaga	tgctgaagag	tttggaaagt	ctgaagctga	3420
aaaataccaa	gatcgaagtt	ttggaagagg	agctcagact	ggcccagagt	gccaactcgg	3480
aaaactgtaa	taagaacaaa	ttcctggatc	agaacctgca	gaaataccag	gcagagtgtt	3540
cccagttcaa	agcgaagctt	gcgagcctgg	aggagctgaa	gagacaggct	gagctggatg	3600
ggaagtccgc	taagcaaaat	ctagacaagt	gctacggcca	aataaaaagaa	ctcaatgaga	3660
agatcacccg	actgacttat	gagattgaag	atgaaaagag	aagaagaaaa	tctgtggaag	3720
acagatttga	ccaacagaag	aatgactatg	accaactgca	gaaagcaagg	caatgtgaaa	3780
aggagaacct	tggttggcag	aaattagagt	ctgagaaaagc	catcaaggag	aaggagtacg	3840
agattgaaag	gttgagggtt	ctactgcag	aagaaggcac	ccggaagaga	gaatatgaaa	3900
atgagctggc	aaaggtaaga	aacctcata	atgaggagat	gagtaattta	aggaacaagt	3960
atgaaacaga	gattaacatt	acgaagacca	ccatcaagga	gatatccatg	caaaaagagg	4020
atgattccaa	aatctttaga	aaccagcttg	atagactttc	aagggaataa	cgagatctga	4080
aggatgaaat	tgtcaggctc	aatgacagca	tcttgcaggc	cactgagcag	cgaaggcgag	4140
ctgaagaaaa	cgcccttcag	caaaaaggcct	gtggtcttga	gataatgcag	aagaagcagc	4200
atctggagat	agaactgaag	caggtcatgc	agcagcgctc	tgaggacaat	gcccggcaca	4260
agcagtcctt	ggaggaggct	gccaagacca	ttcaggacaa	aaataaggag	atcgagagac	4320
tcaaagctga	gtttcaggag	gaggccaagc	gccgttggga	atatgaaaat	gaactgagta	4380
aggtaagaaa	caattatgat	gaggagatca	ttagcttaaa	aatcagttt	gagaccgaga	4440
tcaacatcac	caagaccacc	atccaccagc	tcaccatgca	gaagggaagag	gataccagtg	4500
gctaccgggc	tcagatagac	aatctcacc	gagaaaacag	gagcttatct	gaagaaataa	4560
agaggctgaa	gaacactcta	acccagacca	cagagaatct	caggagggtg	gaagaagaca	4620
tccaacagca	aaaggccact	ggctctgagg	tgtctcagag	gaaacagcag	ctggagggtg	4680
agctgagaca	agtcaactcag	atgcgaacag	aggagagcgt	aagatataag	caatctcttg	4740
atgatgctgc	caaaaccatc	caggataaaa	acaaggagat	agaaagggtta	aaacaactga	4800
tcgacaaaaga	aacaaatgac	cggaaatgcc	tggaaagtga	aaacgcgaga	ttacaaaagg	4860
tccagtatga	cctgcagaaa	gcaaacagta	gtgcgacgga	gacaataaac	aaactgaagg	4920
ttcaggagca	agaactgaca	cgectgagga	tcgactatga	aagggtttcc	caggagagga	4980
ctgtgaagga	ccaggatatc	acgcggttcc	agaactctct	gaaagagctg	cagctgcaga	5040
agcagaaggt	ggaagaggag	ctgaatcggc	tgaagaggac	cgcgtcagaa	gactcctgca	5100

agaggaagaa	gctggaggaa	gagctggaag	gcatgaggag	gtcgctgaag	gagcaagcca	5160
tcaaaatcac	caacctgacc	cagcagctgg	agcaggcatc	cattgttaag	aagaggagt	5220
aggatgacct	ccggcagcag	agggacgtgc	tggatggcca	cctgaggga	aagcagagga	5280
cccaggaaga	gctgaggagg	ctctcttctg	aggtcgaggc	cctgaggcgg	cagttactcc	5340
aggaacagga	aagtgtcaaa	caagctcact	tgaggaatga	gcatttccag	aaggcgatag	5400
aagataaaa	cagaagctta	aatgaaagca	aaatagaaat	tgagaggctg	cagtctctca	5460
cagagaacct	gaccaaggag	cacttgatgt	tagaagaaga	actgcggaac	ctgaggctgg	5520
agtacgatga	cctgaggaga	ggacgaagcg	aagcggacag	tgataaaaat	gcaaccatct	5580
tggaactaag	gagccagctg	cagatcagca	acaaccggac	cctggaactg	caggggctga	5640
ttaatgattt	acagagagag	agggaaaatt	tgagacagga	aattgagaaa	ttccaaaagc	5700
aggcttttaga	ggcatctaata	aggattcagg	aatcaaagaa	tcagtgtact	caggtggtac	5760
aggaaagaga	gagccttctg	gtgaaaatca	aagtcctgga	gcaagacaag	gcaaggctgc	5820
agaggctgga	ggatgagctg	aatcgtgcaa	aatcaactct	agaggcagaa	accagggtga	5880
aacagcgctt	ggagtgtgag	aaacagcaaa	ttcagaatga	cctgaatcag	tggaagactc	5940
aatattccc	caaggaggag	gctattagga	agatagaatc	ggaaagagaa	aagagtgaga	6000
gagagaagaa	cagtcttagg	agtgagatcg	aaagactcca	agcagagatc	aagagaattg	6060
aagagagggtg	caggcgtaag	ctggaggatt	ctaccaggga	gacacagtca	cagttagaaa	6120
cagaacgctc	ccgatatcag	agggagattg	ataaactcag	acagcgccca	tatgggtccc	6180
atcgagagac	ccagactgag	tgtgagtggg	ccgttgacac	ctccaagctg	gtgtttgatg	6240
ggctgaggaa	gaagggtgaca	gcaatgcagc	tctatgagtg	tcagctgata	gacaaaacaa	6300
ccttggaaca	actattgaag	gggaagaagt	cagtgggaaga	agttgcttct	gaaatccagc	6360
cattccttctg	gggtgcagga	tctatcgctg	gagcatctgc	ttctcctaag	gaaaaatact	6420
ctttggtaga	ggccaagaga	aagaaattaa	tcagcccaga	atccacagtc	atgcttctgg	6480
aggcccaggc	agctacaggt	ggtataattg	atccccatcg	gaatgagaag	ctgactgtcg	6540
acagtgccat	agctcgggac	ctcattgact	tcgatgaccg	tcagcagata	tatgcagcag	6600
aaaaagctat	cactgggttt	gatgatccat	tttcaggcaa	gacagtatct	gtttcagaag	6660
ccatcaagaa	aaatttgatt	gatagagaaa	ccggaatgcg	cctgctggaa	gccagattg	6720
cttcaggggg	tgtagtagac	cctgtgaaca	gtgtcttttt	gccaaaagat	gtcgccttgg	6780
cccgggggct	gattgataga	gatttgtatc	gatccctgaa	tgatccccga	gatagtcaga	6840
aaaactttgt	ggatccagtc	accaaaaaga	aggtcagtta	cgtgcagctg	aaggaacggt	6900
gcagaatcga	accacatact	ggtctgctct	tgttttcagt	acagaagaga	agcatgtcct	6960
tccaaggaat	cagacaacct	gtgaccgtca	ctgagctagt	agattctggt	atattgagac	7020
cgtccactgt	caatgaactg	gaatctggtc	agatttctta	tgacgagggt	ggtgagagaa	7080
ttaaggactt	cctccagggt	tcaagctgca	tagcaggcat	atacaatgag	accacaaaac	7140
agaagcttgg	catttatgag	gccatgaaaa	tggctttagt	ccgacctggt	actgctctgg	7200
agttgctgga	agcccaagca	gctactggct	ttatagtgga	tctgtttagc	aacttgaggt	7260
taccagtgga	ggaagcctac	aagagaggtc	tgggtgggcat	tgagttcaaa	gagaagctcc	7320
tgtctgcaga	acgagctgtc	actgggtata	atgatcctga	aacaggaaac	atcatctctt	7380
tgttccaagc	catgaataag	gaactcatcg	aaaaggggcca	cggatttcgc	ttattagaag	7440
cacagatcgc	aaccgggggg	atcattgacc	caaaggagag	ccatcgttta	ccagttgaca	7500
tagcatataa	gaggggctat	ttcaatgagg	aactcagtga	gattctctca	gatccaagt	7560
atgataccaa	aggatttttt	gaccccaaca	ctgaagaaaa	tcttacctat	ctgcaactaa	7620
aagaaagatg	cattaaggat	gaggaaacag	ggctctgtct	tctgcctctg	aaagaaaaga	7680
agaaacaggt	gcagacatca	caaaagaata	ccctcaggaa	gcgtagagt	gtcatagttg	7740
accagaaac	caataaagaa	atgtctgttc	aggaggccta	caagaagggc	ctaattgatt	7800
atgaaacctt	caaagaactg	tgtgagcagg	aatgtgaatg	ggaagaaata	accatcacgg	7860
gatcagatgg	ctccaccagg	gtggtcctgg	tagatagaaa	gacaggcagt	cagtatgata	7920
ttcaagatgc	tattgacaag	ggccttggtg	acaggaaggt	ctttgatcag	taccgatccg	7980
gcagcctcag	cctcactcaa	tttgctgaca	tgatctcctt	gaaaaatgg	gtcggcacca	8040
gcagcagcat	gggcagtgg	gtcagcgatg	atgttttttag	cagctcccga	catgaatcag	8100
taagtaagat	ttccaccata	tccagcgtca	ggaatttaac	cataaggagc	agctcttttt	8160
cagacacctt	ggaagaatcg	agccccattg	cagccattctt	tgacacagaa	aacctggaga	8220
aaatctccat	tacagaaggt	atagagcggg	gcactgttga	cagcatcacg	ggtcagaggc	8280
ttctggaggc	tcaggcctgc	acaggtggca	tcattccacc	aaccacgggc	cagaagctgt	8340

```

cacttcagga cgcagtctcc caggggtgtga ttgaccaaga catggccacc agcgtgaagc 8400
ctgctcagaa agccttcata ggcttcgagg gtgtgaaggg aaagaagaag atgtcagcag 8460
cagaggcagt gaaagaaaaa tggctcccgt atgaggctgg ccagcgcttc ctggagttcc 8520
agtacctcac gggaggtctt gttgaccggg aagtgcattg gaggataagc accgaagaag 8580
ccatccggaa ggggttcata gatggccggc ccgcacagag gctgcaagac accagcagct 8640
atgccaaaat cctgacctgc cccaaaacca aattaaaaat atcctataag gatgccataa 8700
atcgctccat ggtagaagat atcactgggc tgcgccttct ggaagccgcc tccgtgtcgt 8760
ccaagggctt acccagccct tacaacatgt ctccggctcc ggggtcccgc tccggctccc 8820
gctcgggatc tcgctccgga tctcgctccg ggtcccgcag tgggtcccgg agaggaagct 8880
ttgacgccac aggggaattct tctactctt attcctactc atttagcagt agttctattg 8940
ggcactag

```

```

<210> 120
<211> 587
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(587)
<223> n = A,T,C or G

```

```

<400> 120
cgtcctaagc acttagacta catcagggaa gaacacagac cacatccctg tctcatgcg 60
gcttatgttt tctggaagaa agtggagacc nagtccttgg ctttagggct ccccggtgg 120
gggctgtgca ntccggtcag ggcgggaagg gaaatgcacc gctgcatgtg aacttacagc 180
ccaggcggat gcccttccc ttagcactac ctggcctcct gcatccctc gcctcatgtt 240
cctcccacct tcaanaaatg aanaacccca tgggcccgag cccttgccct ggggaaccaa 300
ggcagccttc caaaactcag gggctgaagc anactattag ggcaggggct gactttgggt 360
gacactgccc attccctctc agggcagctc angtcacccn ggnetcttga acccagcctg 420
ttcctttgaa aaagggcaaa actgaaaagg gcttttctta naaaagaaa aaccagggaa 480
ctttgccagg gcttcnntnt taccaaaacn ncttctcnng gatttttaat tccccattng 540
gcctccactt accnggggcn atgcccacaa attaanaatt tcccatc 587

```

```

<210> 121
<211> 619
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(619)
<223> n = A,T,C or G

```

```

<400> 121
cactagtagg atagaaacac tgtgtcccgga gagtaaggag agaagctact attgattaga 60
gcctaacca ggttaactgc aagaagagggc gggatacttt cagctttcca tgtaactgta 120
tgcataaagc caatgtagtc cagttttctaa gatcatgttc caagctaact gaatccact 180
tcaatacaca ctcatgaact cctgatggaa caataacagg cccaagcctg tggatgatg 240
tgcacacttg ctgactcan aaaaaatact actctcataa atgggtggga gtattttggt 300
gacaacctac tttgcttggc tgagtgaagg aatgatattc atatattcat ttattccatg 360
gacatttagt tagtgctttt tatataccag gcatgatgct gagtgcact cttgtgtata 420
tttccaaatt tttgtacagt cgctgcacat atttgaaatc atatattaag acttccaaaa 480
aatgaagtcc ctggtttttc atggcaactt gatcagtaaa ggattcncct ctgttttgga 540

```

cttaaaacat ctactatatn gttnanatga aattcctttt ccccnctcc cgaaaaaana 600
aagtggtagg gaaaaaaa 619

<210> 122
<211> 1475
<212> DNA
<213> Homo sapien

<400> 122

tccacctgtc	cccgccagcg	cggtctcgcg	cctcctgccg	cagccaccga	gccgcogtct	60
agcgccccga	cctcgccacc	atgagagccc	tgctggcgcg	cctgcttctc	tgcgtcctgg	120
tcgtgagcga	ctccaaaggc	agcaatgaac	ttcatcaagt	tccatcgaac	tgtgactgtc	180
taaatggagg	aacatgtgtg	tccaacaagt	acttctccaa	cattcactgg	tgaactgcc	240
caaagaaatt	cggagggcag	caactgtgaa	tagataagtc	aaaaacctgc	tatgagggga	300
atggtcactt	ttaccgagga	aaggccagca	ctgacaccat	gggcccggcc	tgcctgccct	360
ggaactctgc	caactgtcct	cagcaaacgt	accatgccc	cagatctgat	gctcttcage	420
tgggctggg	gaaacataat	tactgcagga	acccagacaa	ccggaggcga	ccctggtgct	480
atgtgcaggt	gggcctaaag	ccgcttgctc	aagagtgc	ggtgcatgac	tgcgcagatg	540
gaaaaaagcc	ctcctctcct	ccagaagaat	taaaatttca	gtgtggccaa	aagactctga	600
ggccccgctt	taagattatt	gggggagaat	tcaccaccat	cgagaaccag	ccctgggttg	660
cggccatcta	caggaggcac	cgggggggct	ctgtcaccta	cgtgtgtgga	ggcagcctca	720
tcagcccttg	ctgggtgatc	agcgccacac	actgcttcat	tgattaccca	aagaaggagg	780
actacatcgt	ctacctgggt	cgtcgaaggc	ttaactccaa	cacgcaaggg	gagatgaagt	840
ttgaggtgga	aaacctcatc	ctacacaagg	actacagcgc	tgacacgctt	gtcaccaca	900
acgacattgc	cttgctgaag	atccgttcca	aggagggcag	gtgtgcgcag	ccatcccga	960
ctatacagac	catctgcctg	ccctcgatgt	ataacgatcc	ccagtttggc	acaagctgtg	1020
agatcactgg	ctttggaaaa	gagaattcta	ccgactatct	ctatccggag	cagctgaaga	1080
tgactgttgt	gaagctgatt	tcaccaccgg	agtgtcagca	gccccactac	tacggctctg	1140
aagtcaccac	caaaatgctg	tgtgctgctg	acccacagtg	gaaaacagat	tcctgccagg	1200
gagactcagg	gggacccctc	gtctgttccc	tccaaggccg	catgactttg	actggaattg	1260
tgagctgggg	ccgtggatgt	gcccgaagg	acaagccagg	cgtctacacg	agagtctcac	1320
acttcttacc	ctggatccgc	agtcacacca	aggaagagaa	tggcctggcc	ctctgagggg	1380
cccaggggag	gaaacgggca	ccaccgcgtt	tcttgctggg	tgtcattttt	gcagtagagt	1440
catctccatc	agctgtaaga	agagactggg	aagat			1475

<210> 123
<211> 2294
<212> DNA
<213> Homo sapien

<400> 123

cagcgccggc	tcgcgccttc	ctgccgcagc	caccgagccg	ccgtctagcg	ccccgacctc	60
gccaccatga	gagccctgct	ggcgccgctg	cttctctgcg	tcctggtcgt	gagcgactcc	120
aaaggcagca	atgaacttca	tcaagttcca	tcgaactgtg	actgtctaaa	tggaggaaca	180
tgtgtgtcca	acaagtactt	ctccaacatt	caactggtgc	actgcccata	gaaattcgga	240
gggcagcact	gtgaaataga	taagtcaaaa	acctgctatg	aggggaatgg	tcacttttac	300
cgaggaaagg	ccagcactga	caccatgggc	cggccctgcc	tgccctggaa	ctctgccact	360
gtccttcagc	aaacgtacca	tgcccacaga	tctgatgctc	ttcagctggg	cctggggaaa	420
cataattact	gcaggaaccc	agacaaccgg	aggcgacct	ggtgctatgt	gcaggtgggc	480
ctaaagccgc	ttgtccaaga	gtgcatgggt	catgactgcg	cagatggaaa	aaagccctcc	540
tctcctccag	aagaattaaa	atttcagtgt	ggccaaaaga	ctctgaggcc	ccgctttaag	600
attattgggg	gagaattcac	caccatcgag	aaccagccct	ggtttgccgc	catctacagg	660
aggcaccggg	ggggctctgt	cacctacgtg	tgtggaggca	gcctcatcag	cccttgctgg	720
gtgatcagcg	ccacacactg	cttcattgat	tacccaaaaga	aggaggacta	catcgtctac	780

ctgggtcgct	caaggcttaa	ctccaacacg	caaggggaga	tgaagtttga	ggtggaaaac	840
ctaatectac	acaaggacta	cagcgctgac	acgcttgctc	accacaacga	cattgccttg	900
ctgaagatcc	gttccaagga	gggcaggtgt	ggcagccat	cccggactat	acagaccatc	960
tgctgccc	cgatgtataa	cgatccccag	tttggcacaa	gctgtgagat	cactggcttt	1020
ggaaaagaga	attctaccga	ctatctctat	ccggagcagc	tgaaaatgac	tgttggaag	1080
ctgatttccc	accgggagtg	tcagcagccc	caactactacg	gctctgaagt	caccacaaa	1140
atgctgtgtg	ctgctgaccc	acagtggaaa	acagattcct	gccagggaga	ctcaggggga	1200
cccctcgtct	gttccctcca	aggccgcag	actttgactg	gaattgtgag	ctggggccgt	1260
ggatgtgccc	tgaaggacaa	gccaggcgct	tacacgagag	tctcacactt	cttaccctgg	1320
atccgcagtc	acaccaagga	agagaatggc	ctggccctct	gaggggtccc	agggaggaaa	1380
cgggcaccac	ccgctttctt	gctggttgct	attttgagct	agagtcactt	ccatcagctg	1440
taagaagagc	tgggaatata	ggctctgcac	agatggattt	gcctgtgcca	ccaccagggc	1500
gaacgacaat	agctttaccc	tcaggcatag	gcctgggtgc	tggctgccc	gacccctctg	1560
gccaggatgg	aggggtggtc	ctgactcaac	atgttactga	ccagcaactt	gtctttttct	1620
ggactgaagc	ctgcaggagt	taaaaagggc	agggcatctc	ctgtgcatgg	gctcgaagg	1680
agagccagct	cccccgaccg	gtgggcattt	gtgaggccca	tggttgagaa	atgaataatt	1740
tcccaattag	gaagtgtaa	cagctgaggt	ctcttgagg	agcttagcca	atgtgggagc	1800
agcggtttgg	ggagcagaga	cactaacgac	ttcagggcag	ggctctgata	ttccatgaat	1860
gtatcaggaa	atatatatgt	gtgtgtatgt	ttgcacactt	gtgtgtgggc	tgtgagtgt	1920
agtgtgagta	agagctggtg	tctgattggt	aagtcataat	atttccttaa	actgtgtgga	1980
ctgtgatgcc	acacagagtg	gtctttctgg	agaggttata	ggtcactcct	ggggcctctt	2040
gggtccccc	cgtgacagtg	cctgggaatg	tattattctg	cagcatgacc	tgtgaccagc	2100
actgtctcag	tttcactttc	acatagatgt	ccctttcttg	gccagttatc	ccttcctttt	2160
agcctagtgc	atccaatcct	cactgggtgg	ggtgaggacc	actcctgtac	actgaatatt	2220
tatatttcac	tatttttatt	tatatttttg	taatttttaa	taaaagtgat	caataaaatg	2280
tgatttttct	gatg					2294

<210> 124

<211> 956

<212> DNA

<213> Homo sapien

<400> 124

gatgagttcc	gcaccaagtt	tgagacagac	caggccctgc	gcctgagtgt	ggaggccgac	60
atcaatggcc	tgcgcaggg	gctggatgag	ctgaccctgg	ccagagccga	cctggagatg	120
cagattgaga	acctcaagga	ggagctggcc	tacctgaaga	agaaccacga	ggaggagatg	180
aacgccctgc	gaggccaggt	gggtggtgag	atcaatgtgg	agatggacgc	tgccccaggc	240
gtggacctga	gccgcatact	caacgagatg	cgtgaccagt	atgagaagat	ggcagagaag	300
aaccgcaagg	atgccgagga	ttggttcttc	agcaagacag	aggaaactgaa	ccgcgaggtg	360
gccaccaaca	gtgagctgg	gcagagtggc	aagagtgaga	tctcggagct	ccggcgcacc	420
atgcaggcct	tggagataga	gctgcagtc	cagctcagca	tgaaagcatc	cctggagggc	480
aacctggcgg	agacagagaa	ccgctactgc	gtgcagctgt	cccagatcca	ggggtgatt	540
ggcagcgtgg	aggagcagct	ggcccagctt	cgctgcgaga	tggagcagca	gaaccaggaa	600
tacaaaatcc	tgctggatgt	gaagacgcgg	ctggagcagg	agattgccac	ctaccgccgc	660
ctgctggagg	gagaggatgc	ccacctgact	cagtacaaga	aagaaccgg	gaccacccgt	720
caggtgcgta	ccattgtgga	agaggtccag	gatggcaagg	tcatctctc	ccgcgagcag	780
gtccaccaga	ccaccgcgtg	aggactcagc	taccccgcc	ggccacccag	gaggcagggg	840
cgcagccgcc	ccatctgccc	cacagtctcc	ggcctctcca	gcctcagccc	cctgcttcag	900
tcccttcccc	atgcttctt	gcctgatgac	aataaaagct	tgttgactca	gctatg	956

<210> 125

<211> 486

<212> DNA

<213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(486)
 <223> n = A,T,C or G

<400> 125

aaattatata	tagtgnttca	gctcccattg	tggtgttcat	agtcttctag	gaacagataa	60
acttaagtat	tcaattcact	cttggcattt	tttctttaat	ataggctttt	tagcctattt	120
ttggaaaact	gcttttcttc	tgagaacctt	attctgaatg	tcatcaactt	taccaaacct	180
tctaagtcca	gagctaactt	agtactgttt	aagttactat	tgactgaatt	ttcttcattt	240
tctgttttagc	cagtgttacc	aaggtaagct	ggggaatgaa	gtataccaac	ttctttcaga	300
gcatttttagg	acattatggc	agcttttagaa	ggctgtcttg	tttctagcca	agggagagcc	360
agcgcagggt	ttggatacta	gagaaaagtc	tttgcttgta	ctattgccat	tttagaaagc	420
tctgatgtga	attcaaattt	tacctctgtt	acttaaagcc	aacaatttta	aggcagtagt	480
tttact						486

<210> 126
 <211> 3552
 <212> DNA
 <213> Homo sapien

<400> 126

cggcaggcag	gtctcgtctc	ggcaccctcc	cggcgcccg	gttctcctgg	ccctgcccg	60
catcccgatg	gccgcgctg	ggccccggcg	ctccgtgcgc	ggagccgtct	gcctgcctct	120
gctgctgacc	ctcgtgatct	tcagtcgtgc	tggtgaagcc	tgcaaaaagg	tgatacttaa	180
tgtaccttct	aaactagagg	cagacaaaat	aattggcaga	gttaatttgg	aagagtgcct	240
caggctctgca	gacctcatcc	ggtcaagtga	tcctgatttc	agagttctaa	atgatgggtc	300
agtgtacaca	gccagggtg	ttgcgctgtc	tgataagaaa	agatcattta	ccatattggc	360
ttctgacaaa	aggaacacaga	cacagaaaga	ggttactgtg	ctgctagaac	atcagaagaa	420
ggtatcgaag	acaagacaca	ctagagaaac	tgttctcagg	cgtgccaaaga	ggagatgggc	480
acctattcct	tgctctatgc	aagagaattc	cttgggacct	ttccatttgt	ttcttcaaca	540
agttgaatct	gatgcagcac	agaactatac	tgtcttctac	tcaataagtg	gacgtggagt	600
tgataaagaa	cctttaaatt	tgttttatat	agaaagagac	actggaaatc	tattttgcac	660
tcggcctgtg	gatcgtgaag	aatatgatgt	ttttgatttg	attgcttatg	cgtcaactgc	720
agatggatat	tcagcagatc	tgccccctcc	actaccatc	agggtagagg	atgaaaatga	780
caaccacct	gttttcacag	aagcaattta	taattttgaa	gttttggaag	gtagtagacc	840
tggtactaca	gtgggggtg	tttgtgccac	agacagagat	gaaccggaca	caatgcatac	900
gcgctgaaa	tacagcattt	tgacgcagac	accaaggtca	cctgggctct	tttctgtgca	960
tcccagcaca	ggcgtaatca	ccacagtctc	tcattatttg	gacagagagg	ttgtagacaa	1020
gtactcattg	ataatgaaag	tacaagacat	ggatggccag	ttttttggat	tgataggcac	1080
atcaacttgt	atcataacag	taacagattc	aatgataat	gcacccactt	tcagacaaaa	1140
tgcttatgaa	gcatttgtag	aggaaaatgc	attcaatgtg	gaaatcttac	gaatacctat	1200
agaagataag	gatttaatta	acactgccaa	ttggagagtc	aattttacca	ttttaagggt	1260
aatgaaaat	ggacatttca	aaatcagcac	agacaaaaga	actaatgaag	gtgttctttc	1320
tgttgtaaag	ccactgaatt	atgaagaaaa	ccgtcaagtg	aacctggaag	ttggagtaaa	1380
caatgaagcg	ccatttgcta	gagatatcc	cagagtgcac	gccttgaaac	gagccttggt	1440
tacagttcat	gtgagggatc	tggtatgagg	gcctgaatgc	actcctgcag	cccaatatgt	1500
gcgattaaa	gaaaacttag	cagtggggtc	aaagatcaac	ggctataagg	catatgaccc	1560
cgaataataga	aatggcaatg	gtttaaggta	caaaaaatct	catgatccta	aaggttggat	1620
caccattgat	gaaatttcag	ggtcaatcat	aacttccaaa	atcctggata	gggaggttga	1680
aactcccaaa	aatgagttgt	ataatattac	agtcctggca	atagacaaag	atgatagatc	1740
atgtactgga	acacttgctg	tgaacattga	agatgtaaat	gataatccac	cagaaatact	1800
tcaagaatat	gtagtcattt	gcaaaccaaa	aatggggtat	accgacattt	tagctgttga	1860

tcttgatgaa	cctgtccatg	gagctccatt	ttatttcagt	ttgcccata	cttctccaga	1920
aatcagtaga	ctgtggagcc	tcaccaaagt	taatgatata	gctgcccgtc	tttcatatca	1980
gaaaaatgct	ggatttcoaag	aatataccat	tcctattact	gtaaaagaca	gggccggcca	2040
agctgcaaca	aaattattga	gagttaatct	gtgtgaatgt	actcatccaa	ctcagtgctg	2100
tgcgacttca	aggagtacag	gagtaatact	tggaaaatgg	gcaatccttg	caatattact	2160
gggtatagca	ctgctctttt	ctgtattgct	aacttttagta	tgtggagitt	ttgggtgcaac	2220
taaagggaaa	cgttttctcg	aagatttagc	acagcaaaac	ttaattatat	caaacacaga	2280
agcacctgga	gacgatagag	tgtgctctgc	caatggattt	atgacccaaa	ctaccaacaa	2340
ctctagccaa	ggtttttctg	gtactatggg	atcaggaatg	aaaaatggag	ggcaggaaac	2400
cattgaaatg	atgaaaggag	gaaaccagac	cttggaatcc	tgcggggggg	ctgggcatca	2460
tcataccctg	gactcctgca	ggggaggaca	cacggagggtg	gacaactgca	gatacactta	2520
ctcggagtg	cacagtttta	ctcaaccccc	tctcggtgaa	aaattgcatc	gatgtaatca	2580
gaatgaagac	cgcattgccat	cccaagatta	tgtcctcact	tataactatg	agggaagagg	2640
atctccagct	ggttctgtgg	gctgctgcag	tgaagagcag	gaagaagatg	gccttgactt	2700
tttaaataat	ttggaaccca	aattttattac	attagcagaa	gcatgcacaa	agagataatg	2760
tcacagtgt	acaattaggt	ctttgtcaga	cattctggag	gtttccaaaa	ataatattgt	2820
aaagttcaat	ttcaacatgt	atgtatatga	tgattttttt	ctcaattttg	aattatgcta	2880
ctcaccaatt	tatatattta	aagcaagttg	ttgcttatct	tttccaaaaa	gtgaaaaatg	2940
ttaaaacaga	caactggtaa	atctcaaaact	ccagcactgg	aattaagggtc	tctaaagcat	3000
ctgctctttt	ttttttttac	agatatttta	gtaataaata	tgctggataa	atattagtcc	3060
aacaatagct	aagttatgct	aatatcacat	tattatgtat	tcactttaag	tgatagttta	3120
aaaaataaac	aagaaatatt	gagtatcact	atgtgaagaa	agttttggaa	aagaaacaat	3180
gaagactgaa	ttaaattaaa	aatgttgcag	ctcataaaga	attggactca	cccctactgc	3240
actaccaaat	tcatttgact	ttggaggcaa	aatgtgttga	agtgccttat	gaagtagcaa	3300
ttttctatag	gaatatagtt	ggaaataaat	gtgtgtgtgt	atattattat	taatcaatgc	3360
aatattttaa	tgaaatgaga	acaaagagga	aaatggtaaa	aacttgaaat	gaggctgggg	3420
tatagtttgt	cctacaatag	aaaaaagaga	gagcttccta	ggcctgggct	cttaaattgct	3480
gcattataac	tgagtctatg	aggaaatagt	tcctgtccaa	tttgtgtaat	ttgtttaaaa	3540
ttgtaataaa	at					3552

<210> 127

<211> 754

<212> DNA

<213> Homo sapien

<400> 127

tttttttttt	ttgtcattgt	tcattgattt	taatgagaaa	gctaagagag	gaaataagta	60
gcctttcaaa	ggtcacacag	aagtaagtga	cagatccagg	attcatatcc	aagcattctg	120
gctctagtgt	ccatgcttct	caaccattat	gacccaatat	tcaacccaaat	caatactgaa	180
ggacacgtga	aatgtatccg	gtatttttact	attacaaaca	aaaatccaat	gaacattctt	240
gaagacatac	acaaaaataa	tggttacaat	agaagttact	ggaattgaaa	ttttggttca	300
acctatatta	aaatgtaagg	cttttgatat	agctaataga	tttttgaaat	gatcagtcct	360
aacgtttgta	ggggagcaca	ctcctgcatg	gggaaaagat	tcactgtgaa	gcacagagca	420
cctttatggg	tggatcatct	tgtcattaaa	gttcaggcgt	tatctatcct	gtaagtggca	480
gaatcaagac	tgcaatatcg	cctgcttttc	tttttaactc	atgttttccc	ttgactacac	540
tggtcctcaa	agtaaaaccc	ctgtgtcagt	gtactattca	tggaaatactc	tgcaattata	600
accaccttct	aatactttta	atacccaatc	aaaattttatt	atacatatgt	atcatagata	660
ctcatctgta	aagctgtgct	tcaaaatagt	gatctcttcc	caacattaca	atatatatta	720
atgatgtcga	acctgcccgg	gcggccgctc	gaag			754

<210> 128

<211> 374

<212> DNA

<213> Homo sapien

<400> 128

agggtttgat	taaaaaggca	aatgatttta	ttgttcgata	atctttttaa	aaaataagag	60
gaaggagtaa	aattaaagat	gaaagatgat	ttttatttcc	ttgtgacctc	tatatcccc	120
ttccccctgcc	cttggttaagt	aactcttgat	ggagaaagga	ttaaagactc	ttattttaacc	180
aaaaaacaga	gccagctaata	catttccaaa	ggtagtatc	tccctgctga	cctcttcttt	240
ggtttaattg	aataaaacta	tatgttcata	tatgtattaa	aacaactcag	aataacatct	300
tttcttcctt	agttaaggca	ttataagggc	tatactatca	tccataataa	ccaaggcaat	360
aacttaaaaa	gctg					374

<210> 129

<211> 546

<212> DNA

<213> Homo sapien

<400> 129

agtgtgatgg	atatctgcag	aattcgggct	aagcgtggtc	gcggccccgag	gtctggaact	60
tcccagcacy	tgaaaaggag	cctcctgagc	tgactcggct	aaagccccac	tttcgctcct	120
cctcatttct	gcctactgat	ttccttggag	cattcatctg	aatattaccg	tttgctgtgt	180
aacctggtac	atacatagca	tgactccctg	gaatagagtg	ggctgggggtg	cttatgctgg	240
gagagtgatt	gacatgcact	ttcaagctat	atctaccatt	tgacgcaaag	gagaaaaaat	300
acctcgagta	aattccatca	ttttttataa	catcagcacc	tgctccatca	tcaaggagtc	360
tcagcgtaac	aggatctcca	gtctctggct	caactgtggc	agtgcagtg	gcattaagaa	420
tgggataaaa	tccctgtttc	acattggcat	aatcatcac	aggatgagga	aaatggaggc	480
tgtctctttc	cacaaaggct	tccacagtgg	ctggggggcac	agacctgccc	gggcggccgc	540
tcgaaa						546

<210> 130

<211> 5156

<212> DNA

<213> Homo sapien

<400> 130

accaaccgag	gcgcggggca	gcgaccctg	cagcggagac	agagactgag	cgccccggca	60
ccgcatgcc	tgcgctctgg	ctgggctgct	gcctctgctt	gtcgtcctc	ctgcccgcag	120
cccgggccac	ctccaggagg	gaagtctgtg	attgcaatgg	gaagtccagg	cagtgtatct	180
ttgatcggga	acttcacaga	caaactggta	atggattccg	ctgcctcaac	tgcaatgaca	240
aaactgatgg	cattcactgc	gagaagtgca	agaatggctt	ttaccggcac	agagaaaggg	300
accgctgttt	gccctgcaat	tgtaactcca	aaggttctct	tagtgctcga	tgtgacaact	360
ccggacggtg	cagctgtaaa	ccagggtgtga	caggagccag	atgcgaccga	tgtctgccag	420
gcttccacat	gtcacggat	gcggggtgca	cccaagacca	gagactgcta	gactccaagt	480
gtgactgtga	cccagctggc	atcgagggc	cctgtgacgc	gggccgctgt	gtctgcaagc	540
cagctgtcac	tggagaacgc	tgtgataggt	gtcgatcagg	ttactataat	ctggatgggg	600
ggaaccctga	gggctgtacc	cagtgtttct	gctatgggca	ttcagccagc	tgccgcagct	660
ctgcagaata	cagtgtccat	aagatcacct	ctacctttca	tcaagatgtt	gatggctgga	720
aggctgtcca	acgaaatggg	tctcctgcaa	agctccaatg	gtcacagcgc	catcaagatg	780
tgtttagctc	agcccaacga	ctagaccctg	tctattttgt	ggctcctgcc	aaatttcttg	840
ggaatcaaca	ggtgagctat	ggtcaaagcc	tgtcctttga	ctaccgtgtg	gacagaggag	900
gcagacaccc	atctgcccac	gatgtgattc	tgggaaggtg	tggtctacgg	atcacagctc	960
ccttgatgcc	acttggcaag	aaactgcctt	gtgggctcac	caagacttac	acattcaggt	1020
taaataagca	tccaagcaat	aattggagcc	cccagctgag	ttactttgag	tatcgaaggt	1080
tactgcggaa	tctcacagcc	ctccgcaccc	gagctacata	tggagaatac	agtactgggt	1140
acattgacaa	tgtgaccctg	atttcagccc	gcctgtctc	tggagcccca	gcaccctggg	1200
ttgaacagtg	tatatgtcct	gttgggtaca	aggggcaatt	ctgccaggat	tgtgcttctg	1260

gctacaagag	agattcagcg	agactggggc	cttttggcac	ctgtattcct	tgtaactgtc	1320
aagggggagg	ggcctgtgat	ccagacacag	gagattgtta	ttcaggggat	gagaatcctg	1380
acattgagtg	tgctgactgc	ccaattgggt	tctacaacga	tccgcacgac	ccccgcagct	1440
gcaagccatg	tcctgtgcat	aacgggttca	gctgctcagt	gatgccggag	acggaggagg	1500
tggtgtgcaa	taactgccct	cccgggggtc	ccggtgcccg	ctgtgagctc	tgtgctgatg	1560
gctactttgg	ggaccccttt	ggtgaacatg	gcccagtgag	gccttgtcag	ccctgtcaat	1620
gcaacaacaa	tgtggacccc	agtgcctctg	ggaatttgtg	ccggtcgaca	ggcaggtgtt	1680
tgaagtgtat	ccacaacaca	gccggcatct	actgcgacca	gtgcaaagca	ggctacttcg	1740
gggacccatt	ggctcccaac	ccagcagaca	agtgtcgagc	ttgcaactgt	aaccccatgg	1800
gctcagagcc	tgtaggatgt	cgaagtgtat	gcacctgtgt	ttgcaagcca	ggatttgggtg	1860
gccccaaactg	tgagcatgga	gcattcagct	gtccagcttg	ctataatcaa	gtgaagattc	1920
agatggatca	gtttatgcag	cagcttcaga	gaatggaggc	cctgatttca	aaggctcagg	1980
gtggtgatgg	agtagtacct	gatacagagc	tggaaggcag	gatgcagcag	gctgagcagg	2040
cccttcagga	cattctgaga	gatgccccga	tttcagaagg	tgctagcaga	tcctttgggtc	2100
tcagttggc	caaggtgagg	agccaagaga	acagctacca	gagccgcctg	gatgacctca	2160
agatgactgt	ggaaagagtt	cgggctctgg	gaagtcagta	ccagaaccga	gttcgggata	2220
ctcacaggct	catcactcag	atgcagctga	gcttggcaga	aagtgaagct	tccttgggaa	2280
acactaacat	tcctgcctca	gaccactacg	tggggccaaa	tggctttaaa	agctctggctc	2340
aggaggccac	aagattagca	gaaagccacg	ttgagtcagc	cagtaacatg	gagcaactga	2400
caagggaaac	tgaggactat	tccaaacaag	ccctctcact	ggtgcgcaag	gccctgcatg	2460
aaggagtcgg	aagcgggaagc	ggtagcccgg	acggtgctgt	ggtgcaaggg	cttgtggaaa	2520
aattggagaa	aaccaagtcc	ctggcccagc	agttgacaag	ggaggccact	caagcggaaa	2580
ttgaagcaga	taggtcttat	cagcacagtc	tcgcctcctc	ggattcagtg	tctcggcttc	2640
agggagtcag	tgatcagtc	tttcaggtgg	aagaagcaaa	gaggatcaaa	caaaaagcgg	2700
attcactctc	aagcctggta	accaggcata	tggatgagtt	caagcgtaca	cagaagaatc	2760
tgggaaactg	gaaagaagaa	gcacagcagc	tcttacagaa	tggaaaaagt	gggagagaga	2820
aatcagatca	gctgctttcc	cgtgccaatc	ttgctaaaag	cagagcacaa	gaagcactga	2880
gtatgggcaa	tgccactttt	tatgaagttg	agagcatcct	taaaaacctc	agagagtttg	2940
acctgcaggt	ggacaacaga	aaagcagaag	ctgaagaagc	catgaagaga	ctctcctaca	3000
tcagccagaa	ggttttcagat	gccagtgaca	agaccagca	agcagaaaaga	gccctgggga	3060
gcgctgctgc	tgatgcacag	agggcaaaga	atggggcccg	ggaggccctg	gaaatctcca	3120
gtgagattga	acagagattt	gggagtcga	acttggaaagc	caatgtgaca	gcagatggag	3180
ccttggccat	ggaaaaggga	ctggcctctc	tgaagagtga	gatgagggga	gtggaaggag	3240
agctggaaag	gaaggagctg	gagtttgaca	cgaatatgga	tgcagtagag	atggtgatta	3300
cagaagccca	gaaggttgat	accagagcca	agaacgctgg	ggttacaatc	caagcacac	3360
tcaacacatt	agacggcctc	ctgcatctga	tggaccagcc	tctcagtgta	gatgaagagg	3420
ggctggtctt	actggagcag	aagctttccc	gagccaagac	ccagatcaac	agccaactgc	3480
ggcccatgat	gtcagagctg	gaagagaggg	cacgtcagca	gaggggccac	ctccatttgc	3540
tggagacaag	catagatggg	attctggctg	atgtgaagaa	cttggagaa	attagggaca	3600
acctgcccc	aggctgctac	aatacccagg	ctcttgagca	acagtgaagc	tgccataaat	3660
atttctcaac	tgaggttctt	gggatacaga	tctcagggct	cgggagccat	gtcatgtgag	3720
tgggtgggat	ggggacattt	gaacatgttt	aatgggtatg	ctcaggtcaa	ctgacctgac	3780
cccatctctg	atcccatggc	caggtgggtg	tcttatttga	ccatactcct	tgcttctctga	3840
tgctgggcaa	tgaggcagat	agcactgggt	gtgagaatga	tcaaggatct	ggaccccaaa	3900
gaatagactg	gatggaaaga	caaactgcac	aggcagatgt	ttgcctcata	atagtcgtaa	3960
gtggagtctt	ggaatttggg	caagtgtctg	tgggatatag	tcaacttatt	ctttgagtaa	4020
tgtgactaaa	ggaaaaaact	ttgactttgc	ccaggcatga	aattcttctt	aatgtcagaa	4080
caaggtgcaa	cccagtcaca	ctgtggccag	taaaatacta	ttgcctcata	ttgtcctctg	4140
caagcttctt	gctgatcaga	gttcctctca	cttacaaccc	aggggtgtgaa	catgttcttc	4200
attttcaagc	tggaaagaagt	gagcagtggt	ggagtgagga	cctgtaaggc	aggcccatcc	4260
agagctatgg	tgcttgctgg	tgccctgccac	cttcaagttc	tggacctggg	catgacatcc	4320
tttcttttaa	tgatgccatg	gcaacttaga	gattgcattt	ttattaaagc	atttctacc	4380
agcaaagcaa	atgttgggaa	agtattttact	ttttcgggtt	caaagtgata	gaaaagtgtg	4440
gcttgggcat	tgaagagggt	aaaattctct	agattttatta	gtcctaattc	aatcctactt	4500

ttagaacacc	aaaaatgatg	cgcataaatg	tattttatct	tattttctca	atctcctctc	4560
tctttctctc	accataata	agagaatgtt	cctactcaca	cttcagctgg	gtcacatcca	4620
tccctccatt	catccttcca	tccatctttc	catccattac	ctccatccat	ccttccaaca	4680
tatatattt	gagtacctac	tgtgtgccag	gggttggtg	gacagtggg	acatagtctc	4740
tgcctcata	gagttgattg	tctagtggag	aagacaagca	tttttaaaaa	ataaatttaa	4800
acttacaac	tttgtttgtc	acaagtggg	tttattgcaa	taaccgcttg	gtttgcaacc	4860
tctttgtca	acagaacata	tgttgcaaga	ccctcccatg	ggggcacttg	agttttggca	4920
aggctgacag	agctctgggt	tgtgcacatt	tctttgcatt	ccagctgtca	ctctgtgcct	4980
ttctacaact	gattgcaaca	gactgttgag	ttatgataac	accagtggga	attgctggag	5040
gaaccagagg	cacttccacc	ttggctggga	agactatggt	gctgccttgc	ttctgtatct	5100
ccttggtatt	tcttgaaagt	gtttttaaat	aaagaacaat	tgtagaaaa	aaaaaa	5156

<210> 131

<211> 671

<212> DNA

<213> Homo sapien

<400> 131

aggtctggag	ggcccacagc	cggatgtggg	acaccgggaa	aaagtgggtca	tagcacacat	60
ttttgcatcc	cgtttgcagt	gtgttgacga	cgaagtcctc	ttgtctgtca	ccccacactt	120
cctgggcagc	caycacgagg	atcatgactc	ggaaaataaa	gatgactgtg	atccacacct	180
tcccgatgct	ggtggagtgt	ttgttgacac	ccccgatgaa	agtgtgcagc	gtcccccaat	240
ccattgcgct	ggtttatccc	tgagtcctgt	ttccaacgac	tgccagtgtt	tcagacccaa	300
agaatgaggg	caagatccct	ctgcgagggg	ttcagacctc	cttctcctac	cccactggag	360
tgcctagaag	ccaatgggtg	cacagtgatg	atacgaatgt	caatctttgc	tcggtcagtg	420
aggatgtcgc	ctggaatatt	caaattgaat	tacagatgca	tgaagagggc	gtacaagtta	480
gaatttttct	ttcgccatac	agaaattggt	tagccagatc	ttctgtactt	cttttccttc	540
cctgaccctt	cctgctcccc	aggaagggag	gtcagccccg	tttgcaaaac	acaggatgcc	600
cgtgacaccg	gagacaggtc	ttcttcaccg	acaggaagtg	ccttctggtg	cctgcacggt	660
ttaactgcta	t					671

<210> 132

<211> 590

<212> DNA

<213> Homo sapien

<400> 132

ctgaatggaa	aagcttatgg	ctctgtgatg	atattagtga	ccagcggaga	tgataagctt	60
cttggaatt	gcttaccac	tgtgtcagc	agtggttcaa	caattcactc	cattgccctg	120
ggttcactg	cagccccaaa	tctggaggaa	ttatcacgtc	ttacaggagg	tttaaagtgc	180
tttgttccag	atatatcaaa	ctccaatagc	atgattgatg	ctttcagtag	aatttcctct	240
ggaactggag	acattttcca	gcaacatatt	cagcttgaaa	gtacaggtga	aaatgtcaaa	300
cctcaccatc	aattgaaaaa	cacagtgact	gtggataata	ctgtgggcaa	cgacactatg	360
tttctagtta	cgtggcaggc	cagtggctct	cctgagatta	tattatttga	tctgatgga	420
cgaaaatact	acacaaataa	ttttatcacc	aatctaactt	ttcgacagc	tagtctttgg	480
attccaggaa	cagctaagcc	tgggcactgg	acttacacc	tgaacaatac	ccatcattct	540
ctgcaagccc	tgaaagtgac	agtgacctct	cgcgcctcca	actcagacct		590

<210> 133

<211> 581

<212> DNA

<213> Homo sapien

<400> 133

```

aggctcctgtc cgggggcact gagaactccc tctggaattc ttggggggtg ttgggggagag      60
actgtgggcc  tggagataaa acttgtctcc tctaccacca ccctgtaccc tagcctgcac      120
ctgtcctcat  ctctgcaaag ttcagcttcc ttccccaggt ctctgtgcac tctgtcttgg      180
atgctctggg  gagctcatgg gtggaggagt ctccaccaga gggaggctca ggggactggt      240
tggggcaggg  atgaatattt gagggataaa aattgtgtaa gagccaaaga attggtagta      300
gggggagAAC  agagaggagc tgggctatgg gaaatgattt gaataatgga gctgggaata      360
tggctggata  tctggtacta aaaaagggtc tttaagaacc tacttcttaa tctcttcccc      420
aatccaaacc  atagctgtct gtccagtgtc ctcttctctg ctccagctct gccccaggt      480
cctcctagac  tctgtccctg ggctagggca ggggaggagg gagagcaggg ttgggggaga      540
ggctgaggag  agtgtgacat gtggggagag gaccagacct c                                581

```

<210> 134

<211> 4797

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(4797)

<223> n = A,T,C or G

<400> 134

```

cctgggacca aagtgtgcc cagagctgag ggtcctggag ccacatgaga aggtttctcc      60
ctgtgtacct gtgcagcaca gggtaggggtg agtccactca gctgtctagg agaggaccca      120
ggagcagcag agacncgcca agcctttact cataccatat tctgatecctt ttccagcaaa      180
ttgtggctac taatttgccc cctgaagatc aagatggctc tggggatgac tctgacaact      240
tctccggctc aggtgcaggt gaggttgtca tgggggcccc cccaccccaa gacggcaaca      300
ggtcatgcct gggggcagtg gtcaggcagt ctccctgtgtt tactgagcat gtactgagtg      360
caccctgcct gccctgtctc caccagctg gctccaaagg gcaatgctga ggagaggaat      420
ggggtcgtga gctgctgtta aggagagctc atgcttggag gtgaggtgaa ggctgtgagc      480
tccagaaggc cccagggcgc nctgctgcac gcaggctcat attcactagg aatagcttta      540
ctcactaaga aacctctgga accccttcca gaaggttatt tgactcctga gcctctatatt      600
tctcatctgc aaaatgggaa taataccttg acctgataag cttgtggagc tgtaaggcag      660
cacagagcca gctgggggtg agtccctcag ctggaggtgg tcagggaagc ttcacagagg      720
tgtggggact gggggagaga agtccctcag ctggaggtgg tcagggaagc ttcacagagg      780
aggtggctct tgagtggacc tcaggaagag gggtagagga gctaagggaag gaggctgagg      840
tcatccctgg ggaagtgacc tagcggaggc ctgagagctg caaggtagga tatctgttgt      900
tggaagtgtc tgttgttgga agtgggggcc tttttttcag ggaggggtggg gccagagaag      960
tgtgtgccct gggataagta ggataaccac agtagttatg cccctaaggg atgcccaccc      1020
caccctgtg gtacacagaaa agcttttcca ggtggcctag gcacctgtct cgtgggtcca      1080
gagacaggct gcacctgaca cacacaatgg aaggacagct ctccctgtcc attttccaag      1140
gagcttagcc tcagctgcct tgtccaggta ctagcctccc tcatagcctg agcttggcca      1200
gcccagggtg tctggagcct ccccgaccc acccaacaca ctctgcttct ggtcctcccc      1260
acccccccacc tccccaacac actctgcttc tggctcctgca ggtgctttgc aagatatcac      1320
cttgtcacag cagacccctt ccacttgga ggacacgcag ctccctgacg ctattccac      1380
gtctccagaa cccaccggcc tggaggctac agctgcctcc acctccaccc tgcgggctgg      1440
agagggggccc aaggaggagg aggtgtagt cctgccagaa gtggagcctg gcctcaccgc      1500
ccgggagcag gaggccaccc cccgacccag ggagaccaca cagctcccga ccaactcatca      1560
ggcctcaacg accacagcca ccacggccca ggagcccgc acctccacc cccacaggga      1620
catgcagcct ggccaccatg agacctcaac ccctgcagga cccagccaag ctgaccttca      1680
cactccccac acagaggatg gaggtccttc tgccaccgag agggctgctg aggatggagc      1740
ctccagtcag ctcccagcag cagagggtc tggggagcag gtgagtggcc tctgcattcc      1800
ttgggaaatt gagtgggttg gtcctaattg ctggcacttg gcaggcccta cacctgtgcc      1860
ctgcgcgata tegtattcct caccaggaag acagggcaca ggggcgcctt tcccctaccc      1920

```

ccagggcctc	gcagagcagg	acagactaac	tatgagatca	gagcagaagc	acccttaaag	1980
atcaccacaag	agagggctcc	caaaactcaca	atccaaactt	gcagccctcg	tcgaagagtg	2040
aacgttatac	cagtcatttt	atttatagct	tcgtggattt	acgcttacac	taaatagtct	2100
gctattcata	caaaatgtgt	gctttgtatc	actttttgtg	atatccatgc	catgggtccag	2160
ccagggtcog	gagttgatgt	ggcaagaagg	cctggctttc	gggccctgtg	cgatcctggg	2220
ttgggtgcat	ctgagtgggt	gggtggcaaag	atcagggagg	caggagctgc	ttctgggtct	2280
gtagtggagc	tggttgctgc	tgctggcggg	gacctggcca	acccaatctg	cccctgccct	2340
cccacaggac	ttcacctttg	aaacctcggg	ggagaatacg	gctgtagtgg	ccgtggagcc	2400
tgaccgccgg	aaccagtcct	cagtggatca	gggggccacg	ggggcctcac	agggcctcct	2460
ggacaggaaa	gaggtgctgg	gaggtgagtt	ttctttcagg	ggggtagttt	ggggtgaatt	2520
gctgctgtgg	ggtcaggggtg	gggctgacca	cagccaaggc	cactgctttg	ggaggggtctg	2580
cacgagagcc	caaggagccg	ctgagctgag	ctggccccgt	ctacctgccc	taggggtcat	2640
tgcgggagge	ctcgtggggc	tcactctttgc	tgtgtgcctg	gtgggtttca	tgctgtaccg	2700
catgaagaag	aaggacgaag	gcagctactc	cttgaggagg	cggaaacaag	ccaacggcgg	2760
ggcctaccag	aagcccacca	aacaggagga	attctatgcc	tgacgcggga	gccatgcgcc	2820
cctccgcgcc	tgccactcac	taggccccca	cttgccctct	ccttgaagaa	ctgcaggccc	2880
tgccctcccc	tgccaccagg	ccacctcccc	agcattccag	ccccctctgt	cgctcctgcc	2940
caaggagtcg	tggtgtgtgt	gggagctcca	ctctgcttct	ctgacttctg	cctggagact	3000
tagggcacca	gggttttctc	gcataggacc	tttccaccac	agccagcacc	tggtcatcgca	3060
ccattctgac	tcggtttctc	caaactgaag	cagcctctcc	ccaggtccag	ctctggaggg	3120
gagggggatc	cgactgcttt	ggacctaaat	ggcctcatgt	ggctggaaga	tcctgcgggt	3180
ggggcttggg	gctcacacac	ctgtagcact	tactggtagg	accaagcatc	ttgggggggt	3240
ggccgctgag	tgccagggga	caggagtac	tttgtttcgt	ggggagggtc	aattctagata	3300
tcgacttggt	tttgccacatg	tttccctctag	ttctttgttc	atagcccagt	agaccttggt	3360
acttctgagg	taagttaagt	aagttgatcc	ggtatcccc	catcttgctt	ccctaatacta	3420
tggtcgggag	acagcatcag	ggttaagaag	actttttttt	tttttttttaa	actaggagaa	3480
ccaaatctgg	aagccaaaat	gtaggcttag	tttgtgtgtt	gtctcttgag	tttgtcgtc	3540
atgtgtgcaa	caggggatgg	actatctgtc	tggtggcccc	gttctgggtg	tctgttgcca	3600
ggctggccag	tcagggtgc	cgtggggccg	ccgcctcttt	caagcagtcg	tgccctgtgtc	3660
catgcgctca	gggccatgct	gaggcctggg	cgcctgccac	gttgaggaga	cccgtgtgag	3720
aagtgaatgc	tgggactcag	ccttcagaca	gagaggactg	tagggagggg	ggcagggggc	3780
tgagatcct	cctgcaggct	cacgcccgtc	ctcctgtggc	gcgctctcca	ggggctgctt	3840
cctcctggaa	attgacgagg	ggtgtcttgg	gcagagctgg	ctctgagcgc	ctccatccaa	3900
ggccagggtc	tcggttagct	cctgtggccc	caccctgggc	cctgggctgg	aatcaggaat	3960
attttccaaa	gagtgatagt	cttttgcttt	tggtcaaaact	ctacttaatc	caatgggttt	4020
ttccctgtac	agtagatttt	ccaaatgtaa	taaactttta	tataaagtag	tctgtgaatg	4080
ccactgcctt	cgcttcttgc	ctctgtgctg	tgtgtgacgt	gaccggactt	ttctgcaaac	4140
accaacatgt	tggtgaaact	ggctcgaatc	tctgtgcctt	cgtctttccc	atggggaggg	4200
attctgggtc	cagggctccct	ctgtgtattt	gcttttttgt	tttggctgaa	attctcctgg	4260
aggctggtag	gttcagccaa	ggttttataa	ggctgatgtc	aattttctgt	ttgccaagct	4320
ccaagcccat	cttctaaatg	gcaaagggaag	gtggatggcc	ccagcacagc	ttgacctgag	4380
gctgtgggtc	cagcggaggt	gtggagccga	ggcctacccc	ncagacacct	tgacatcct	4440
cctcccaccc	ggctgcagag	gccaganncc	agcccagggt	cctgcactta	cttgcctatt	4500
tgacaacggt	tcagcgactc	cgttgccac	tcagagagtg	ggccagtcctg	tggtatcagag	4560
atgcaccacc	aagccaaggg	aacctgtgtc	cggatttcga	tactgcgact	ttctgcctgg	4620
agtgtatgac	tgacatgac	tcgggggtgg	ggaaaggggt	cggctgacca	tgctcatctg	4680
ctggtccgtg	ggacggtnc	caagccagag	gtgggttcat	ttgtgtaacg	acaataaacg	4740
gtacttgtca	tttcgggcaa	cggctgctgt	ggtgggtggt	gagtctcttc	ttggcct	4797

<210> 135

<211> 2856

<212> DNA

<213> Homo sapien

<400> 135

tagtcgcggg	tccccgagt	agcacgccag	ggagcaggag	accaaacgac	gggggtcgga	60
gtcagagtcg	cagtgggagt	ccccggaccg	gagcacgagc	ctgagcggga	gagcgccgct	120
cgcacgcccg	tcgccaccgg	cgtacccggc	gcagccagag	ccaccagcgc	agcgtgcca	180
tggagcccag	cagcaagaag	ctgacgggtc	gcctcatgct	ggctgtggga	ggagcagtgc	240
ttggctccct	gcagtttggt	tacaacactg	gagtcataca	tgccccccag	aaggatgatcg	300
aggagttcta	caaccagaca	tgggtccacc	gctatgggga	gagcactctg	cccaccacgc	360
tcaccacgct	ctggtccttc	tcagtggcca	tcttttctgt	tgggggcatg	attggctcct	420
tctctgtggg	ccttttcgtt	aaccgctttg	gccggcggaa	ttcaatgctg	atgatgaacc	480
tgtgtggcct	cgtgtccggc	gtgtcatggt	gcttctcgaa	actgggcaag	tcctttgaga	540
tgtgtatcct	gggcccgttc	atcatcggtg	tgtactgctg	cctgaccaca	ggcttcgtgc	600
ccatgtatgt	gggtgaagtg	tcacccacag	cctttcgtgg	ggccctgggc	accctgcacc	660
agctgggcat	cgtcgtcggc	atcctcatcg	cccaggtgtt	cggcctggac	tcctcatatg	720
gcaacaagga	cctgtggccc	ctgtgtctga	gcatacatct	catcccggcc	ctgtgtcagt	780
gcacgtgtct	gcccttctgc	cccgagagtc	cccgttctct	gctcatcaac	cgcaacgagg	840
agaaccgggc	caagagtgtg	ctaaagaagc	tgcgcgggac	agctgacgtg	acccatgacc	900
tgcaggagat	gaagggaagc	agtcggcaga	tgatgcggga	gaagaaggtc	accatcctgg	960
agctgttccg	ctccccgcgc	taccgccagc	ccatcctcat	cgtgtgtgtg	ctgcagctgt	1020
cccagcagct	gtctggcacc	aacgtgtctt	tctattactc	cacgagcctc	ttcgagaagg	1080
cgggggtgca	gcagcctgtg	tatgccacca	ttggctccgg	tatcgtcaac	acggccttca	1140
ctgtcgtgtc	gctgtttgtg	gtggagcgag	caggccggcg	gaccctgcac	ctcataggcc	1200
tcgctggcat	ggcgggttgt	gccatactca	tgaccatcgc	gctagcactg	ctggagcagc	1260
taccctggat	gtcctatctg	agcatcgtgg	ccatccttgg	ccttgtggcc	ttctttgaag	1320
tgggtcctgg	ccccatccca	tgggtccatg	tggctgaact	cttcagccag	ggtccacgct	1380
cagctgccat	tgcggttgca	ggcttctcca	actggacctc	aaatttcatt	gtgggcatgt	1440
gcttcagta	tgtggagcaa	ctgtgtgggt	cctacgtctt	catcatcttc	actgtgtctc	1500
tgggtctgtt	cttcattctc	acctacttca	aagttcctga	gactaaaggc	cggaccttcg	1560
atgagatcgc	ttccggtctc	cggcaggggg	gagccagcca	aagtgataag	acacccgagg	1620
agctgttcca	tcccctgggg	gctgattccc	aagtgtgagt	cgccccagat	caccagcccc	1680
gctgtctccc	agcagcccta	aggatctctc	aggagcacag	gcagctggat	gagacttcca	1740
aacctgacag	atgtcagccg	agccgggctt	ggggtcctct	tctccagcca	gcaatgatgt	1800
ccagaagaat	attcaggact	taacggctcc	aggattttta	caaaagcaag	actgtttgctc	1860
aaatctattc	agacaagcaa	caggttttat	aattttttta	ttactgattt	tgttatatttt	1920
atatcagcct	gagtctcctg	tgcccacatc	ccaggcttca	ccctgaatgg	ttccatgcct	1980
gaggtgggag	actaagccct	gtcgagacac	ttgccttctt	caccagctca	atctgtaggg	2040
ctggacctat	gtcctaagga	cacactaatc	gaactatgaa	ctacaaagct	tctatcccag	2100
gaggtggcta	tggccaaccg	ttctgtctgg	ctggatctcc	ccactctagg	ggtcaggctc	2160
cattaggatt	tgcctcttcc	catctcttcc	tacccaacca	ctcaaattaa	tctttcttta	2220
cctgagacca	gttgggagca	ctggagtgca	gggaggagag	gggaagggcc	agtctgggct	2280
gccgggttct	agtctccttt	gcactgaggg	ccacactatt	acccatgagaa	gagggcctgt	2340
gggagcctgc	aaactcactg	ctcaagaaga	catggagact	cctgccctgt	tgtgtataga	2400
tgcaagatat	ttatatatat	ttttggttgt	caatattaaa	tacagacact	aagttatagt	2460
atatctggac	aagccaactt	gtaaatacac	cacctcactc	ctgttactta	cctaaacaga	2520
tataaatggc	tggtttttag	aaacatgggt	ttgaaatgct	tgtggattga	gggtaggagg	2580
tttgatggg	agtgtgacag	aagtaagtgg	ggttgcaacc	actgcaacgg	cttagacttc	2640
gactcaggat	ccagtccctt	acacgtacct	ctcatcagtg	tcctcttctg	caaaaatctg	2700
tttgatccct	gttaccacaga	gaatatatac	attctttatc	ttgacattca	aggcatttct	2760
atcacatatt	tgatagttgg	tgttcaaaaa	aacactagtt	ttgtgccagc	cgtgatgtct	2820
aggcttgaaa	tcgcattatt	ttgaatgtga	agggaa			2856

<210> 136

<211> 356

<212> DNA

<213> Homo sapien

<400> 136

```

gggtggagcca aatgaagaaa atgaagatga aagagacaga cacctcagtt tttctggatc      60
aggcattgat gatgatgaag attttatctc cagcaccatt tcaaccacac cacgggcttt      120
tgaccacaca aaacagaacc aggactggac tcagtggaaac ccaagccatt caaatccgga      180
agtgtacttt cagacaacca caaggatgac tgatgtagac agaaatggca ccaactgtta      240
tgaaggaaac tggaaccag aagcacaccc tccccctcatt caccatgagc atcatgagga      300
agaagagacc ccacattcta caagcacaat ccaggcaact cctagtagta caacgg      356

```

<210> 137

<211> 356

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(356)

<223> n = A,T,C or G

<400> 137

```

gcaggtggag aagacatttt attgttcctg ggggtctctgg aggccattg gtggggctgg      60
gtcactggct gcccccgaa cagggcgctg ctccatggct ctgcttgagg tagtctgtgg      120
ctatgtctcc cagcaaggac agaaactcag aaaaatcaat cttcttatcc tcattcttgt      180
cctttttctc aaagacatcg gcgaggtaat ttgtgccctt tttacctcgg cccgcgacca      240
cgctaaggcc aaanttccag acanayggcc gggccgggtnc nataggggan cccaacttgg      300
ggacccaaac tctggcgcgg aaacacangg gcataagctt gnttcctgtg gggaaa      356

```

<210> 138

<211> 353

<212> DNA

<213> Homo sapien

<400> 138

```

aggtccagtc ctccacttgg cctgatgaga gtggggagtg gcaagggacg tttctcctgc      60
aatagacact tagattttctc tcttggtgga agaaaccacc tgtccatcca ctgactcttc      120
tacattgatg tggaaaattgc tgctgctacc accacctcct gaagaggctt cctgatgcc      180
aatgccagcc atcttggcat cctggccctc gagcaggctg cgtaagtag cgatctcctg      240
ctccagccgt gtctttatgt caagcagcat cttgtactcc tggttctgag cctccatctc      300
gcatcggagc tactcagac ctgscgsg mssmcgctam gccgaattcc agc      353

```

<210> 139

<211> 371

<212> DNA

<213> Homo sapien

<400> 139

```

agcgtggtcg cggccgaggt ccateccaag caagattgca gatggcagtg tgaagagaga      60
agacatatcc tacatttcaa agctttgggtg caattcccat cgaccagagt tgggccgacc      120
agccttggaa aggtcactga aaaatcttca attggattat gttgacctct acctatttca      180
ttttccagtg tctgtaaagc caggtgagga agtgatccca aaagatgaaa atggaaaaat      240
actatttgac acagtggatc tctgtgccac gtgggagggc gtggagaagt gtaaagatgc      300
aggattggac ctgcccgggc ggccgctcga aagccgaatt ccagcacact ggcggccggt      360
actagtggat c

```

371

<210> 140
 <211> 370
 <212> DNA
 <213> Homo sapien

<400> 140

tagcgtggtc	gcgcccgagg	tccatctccc	tttgggaact	agggggctgc	tgggtgggaaa	60
tgggagccag	ggcagatggt	gcattccttt	gtgtccctgt	aaatgtggga	ctacaagaag	120
aggagctgcc	tgagtggtag	tttctcttcc	tggtaatcct	ctggcccagc	ctcatggcag	180
aatagaggta	tttttaggct	atttttgtaa	tatggcttct	ggtcaaaatc	cctgtgtagc	240
tgaattccca	agccctgcat	tgtacagccc	cccactcccc	tcaccaccta	ataaagggaat	300
agttaacact	caaaaaaaaa	aaaaaacctg	cccgggcggc	cgctcgaaa	ccgaattcca	360
gcacactggc						370

<210> 141
 <211> 371
 <212> DNA
 <213> Homo sapien

<400> 141

tagcgtggtc	gcgcccgagg	tcctctgtgc	tgccctgtcac	agcccgatgg	taccagcgca	60
gggtgtaggc	agtgcaggag	ccctcatcca	gtggcaggga	acaggggtca	tcactatccc	120
aaggagcttc	agggtcctgg	tactctctcca	cagaatactc	ggagtattca	gagtactcat	180
catctcagg	gggtaccgcg	tcttctctct	ctgcatgaga	gacgcggagc	acaggcacag	240
catggagctg	ggagccggca	gtgtctgcag	cataactagg	gaggggtcgt	gatccagatg	300
cgatgaactg	gccctggcag	gcacagtget	gactcatctc	ttggcgacct	gcccgggcgg	360
ccgctcgaag	c					371

<210> 142
 <211> 343
 <212> DNA
 <213> Homo sapien

<400> 142

gcgttttgag	gccaatggtg	taaaaggaaa	tatcttcaca	taaaaactag	atggaagcat	60
tgtcagaaac	ctctttgtga	tgtttgcttt	caactcacag	agttgaacat	tccttttcat	120
agagcagttt	tgaaacactc	ttttgtagaa	tttgcaagcg	gatgattgga	tcgctatgag	180
gtcttcattg	gaaacgggat	acctttacat	aaaaactaga	cagtagcatt	ctcagaaatt	240
tctttgggat	gtgggcattc	aaccacaga	ggagaacttc	atttgataga	gcagttttga	300
aacacccttt	ttgtagaatc	tacaggtgga	catttagagt	gct		343

<210> 143
 <211> 354
 <212> DNA
 <213> Homo sapien

<400> 143

aggtctgatg	gcagaaaaac	tcagactgtc	tgcaacttta	cagatgggtgc	attggttcag	60
catcaggagt	gggatgggaa	ggaaagcaca	ataacaagaa	aattgaaaga	tgggaaatta	120
gtggtggagt	gtgtcatgaa	caatgtcacc	tgtactcgga	tctatgaaaa	agtagaataa	180
aaattccatc	atcacttttg	acaggagtta	attaagagaa	tgaccaagct	cagttcaatg	240
agcaaatctc	catactgttt	ctttcttttt	tttttcatta	ctgtgttcaa	ttatctttat	300
cataaacatt	ttacatgcag	ctattttcaa	gtgtgttgga	ttaattagga	tcat	354

<210> 144
 <211> 353
 <212> DNA
 <213> Homo sapien

<400> 144

ggtcaaggac	ctgggggacc	cccagggtcca	gcagccacat	gattctgcag	cagacagggg	60
cctagagcac	atctggatct	cagccccacc	cctggcaacc	tgctgccta	gagaactccc	120
aagatgacag	actaagtagg	attctgccat	ttagaataat	tctggatatcc	tgggcggtgc	180
gttaagttgc	ttacttttca	ttctgtctta	cgatagtctt	cagaggtggg	aacagatgaa	240
gaaaccatgc	cccagagaag	gttaagtgc	ttcctcttta	tgagagccagt	gttccaacct	300
aggtttgcct	gataccagac	ctgtggcccc	acctcccatg	caggtctctg	tgg	353

<210> 145
 <211> 371
 <212> DNA
 <213> Homo sapien

<400> 145

caggtctgtc	ataaaactggt	ctggagtttc	tgacgactcc	ttgttcacca	aatgcaccat	60
ttcctgagac	ttgctggcct	ctccgttgag	tccacttggc	tttctgtcct	ccacagctcc	120
attgccactg	ttgatcacta	gctttttctt	ctgccacac	cttcttcgac	tgttgactgc	180
aatgcaaact	gcaagaatca	aagccaaggc	caagagggat	gccaaagatga	tcagccattc	240
tggaatttgg	ggtgtcctta	taggaccaga	ggttgtgttt	gctccacett	cttgactccc	300
atgtgagacc	tcggccgcga	ccacgctaag	ccgaattcca	gcacactggc	ggcccgttac	360
tagtgatcc	g					371

<210> 146
 <211> 355
 <212> DNA
 <213> Homo sapien

<400> 146

ggtcctccgt	cctcttccca	gaggtgtcgg	ggcttgcccc	cagcctccat	cttcgtctct	60
caggatggcg	agtagcagcg	gtcccaaggc	tgaattcatt	gtcggaggga	aatataaact	120
ggtacggaag	atcgggtctg	gtccttcggg	ggacatctat	ttggcgatca	acatcaccaa	180
cggcgaggaa	gtggcagtga	agctagaatc	tcagaaggcc	aggcatcccc	agttgctgta	240
cgagagcaag	ctctataaga	ttcttcaagg	tggggttggc	atccccaca	tacggtggtg	300
tggtcaggaa	aaagactaca	atgtactagt	catggatctt	ctgggacctc	gcctc	355

<210> 147
 <211> 355
 <212> DNA
 <213> Homo sapien

<400> 147

ggtctgttac	aaaatgaaga	cagacaacac	aacatttact	ctgtggagat	atcctactca	60
tactatgcac	gtgctgtgat	tttgaacata	actcgtccca	aaaacttgtc	acgatcatcc	120
tgacttttta	ggttggctga	tccatcaatc	ttgactcaa	ctgttacttc	tttcccagtg	180
ttgttagag	caaagctgac	ctgaacagca	accaatggct	gtagataccc	aacatgcagt	240
tttttcccat	aatatgggaa	atattttaag	tctatcattc	cattatgagg	ataaactgct	300
acatttggtg	tatcttcatt	ctttgaaaca	caatctatcc	ttggcactcc	ttcag	355

<210> 148

<211> 369
 <212> DNA
 <213> Homo sapien

<400> 148
 aggtctctct cccctctctc ctctctctgcc agccaagtga agacatgctt acttcccctt 60
 caccttcctt catgatgtgg gaagagtgtc gcaaccacgc cctagccaac accgcatgag 120
 agggagtgtg ccgagggtct ctgagaaggt ttctctcaca tctagaaaga agcgcttaag 180
 atgtggcagc cctctctctt caagtggctc ttgtctgtt gccctgggag ttctcaaatt 240
 gctgcagcag cctccatcca gcctgaggat gacatcaata cacagaggaa gaagagtcag 300
 gaaaagatga gagaagttac agactctcct gggcgacccc gagagcttac cattcctcag 360
 acttcttca 369

<210> 149
 <211> 620
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(620)
 <223> n = A,T,C or G

<400> 149
 actagtcaaa aatgctaaaa taatttggga gaaaatattt ttttaagtagt gttatagttt 60
 catgtttatc ttttattatg ttttgtgaag ttgtgtcttt tcactaatta cctatactat 120
 gccaatattt cttatatctc atccataaca ttctactac atttgaana naatatgcac 180
 gtgaaactta acactttata aggtaaaaat gaggtttcca anatttaata atctgatcaa 240
 gttcttgtaa ttccaaata gaatggactt ggtctgttaa gggctaagga gaagaggaag 300
 ataagggttaa aagttgttaa tgaccaaaaca ttctaaaaga aatgcaaaaa aaaagtttat 360
 tttcaagcct tcgaactatt taaggaaaagc aaaatcattt cctaaatgca tatcatttgt 420
 gagaatttct cattaatatc ctgaatcatt catttcacta aggtcatgt tnaactccgat 480
 atgtctctaa gaaagtacta ttcatgggtc caaacctggg tgccatantt gggtaaaggc 540
 ttcccttaa gtgtgaaant atttaaatg aaattttcct ctttttaaaa attctttana 600
 aggttaagg gtgttgggga 620

<210> 150
 <211> 371
 <212> DNA
 <213> Homo sapien

<400> 150
 ggtccgatca aaacctgcta cctccccaag actttactag tgccgataaa ctttctcaaa 60
 gagcaaccag taccacttcc ctgtttataa aacctctaac catctctttg ttctttgaac 120
 atgctgaaaa ccacctgggc tgcattgtatg ccgaatttg yaattctttt ctctcaaagt 180
 aaaatttaaa ttttagggatt catttctata ttttcacata ttagtatta ttatttccct 240
 atatgtgtaa ggtgaaattt atggattttg agtgtgcaag aaaatatatt tttaaagctt 300
 tcatttttcc ccagtgaaat gatttagaat tttttatgta aatatacaga atgttttttc 360
 ttacttttat a 371

<210> 151
 <211> 4655
 <212> DNA
 <213> Homo sapien

<400> 151

gggacttgag	ttctgttata	ttcttaagta	gattcatatt	gtaagggctt	cgggggtggg	60
gggttggcaa	aatcctggag	ccagaagaaa	ggacagcagc	attgatcaat	cttacagcta	120
acatgttgta	cctggaaaac	aatgcccgaga	ctcaatttag	tgagccacag	tacacgaacc	180
tggggctcct	gaacagcatg	gaccagcaga	ttcagaacgg	ctcctcgtec	accagtcctt	240
ataacacaga	ccacgcgcag	aacagcgtea	cggegccttc	gccctacgca	cagcccagct	300
ccaccttcga	tgctctctct	coatcaccgc	ccatcccttc	caacaccgac	taccagggcc	360
cgcacagttt	cgacgtgtcc	ttccagcagt	cgagcaccgc	caagtcggcc	acctggacgt	420
attccactga	actgaagaaa	ctctactgcc	aaattgcaaa	gacatgcccc	atccagatca	480
aggtgatgac	cccacctcct	cagggagctg	ttatccgcgc	catgcctgtc	tacaaaaaag	540
ctgagcacgt	cacggagggtg	gtgaagcggt	gccccaacca	tgagctgagc	cgtgaattca	600
acgagggaca	gattgcccct	ytatgtcatt	tgattcgagt	agaggggaac	agccatgccc	660
agtatgtaga	agatcccata	acaggaagac	agagtgtgct	ggtaccttat	gagccacccc	720
aggttggcac	tgaattcacg	acagtcttgt	acaatttcat	gtgtaacagc	agttgtgttg	780
gagggatgaa	cgcgcgtcca	attttaatat	ttgttactct	ggaaaccaga	gatgggcaag	840
tcttggggcg	acgtgtcttt	gaggcccgga	tctgtgcttg	cccaggaaga	gacaggaagg	900
cggatgaaga	tagcatcaga	aagcagcaag	tttcggacag	tacaaagaac	ggtgatggta	960
cgaagcgccc	gtttcgtcag	aacacacatg	gtatccagat	gacatccatc	aagaaacgaa	1020
gatccccaga	tgatgaactg	gtatacttac	cagtgagggg	ccgtgagact	tatgaaatgc	1080
tggtgaagat	caaagagtcc	ctggaactca	tgcatgacct	tcttcagcac	acaattgaaa	1140
cgtacaggca	acagcaacag	cagcagcacc	agcacttact	tcagaaacag	acctcaatac	1200
agtctccatc	ttcatatggt	aacagctccc	cacctctgaa	caaaatgaac	agcatgaaca	1260
agctgccttc	tgtgagccag	cttatcaacc	ctcagcagcg	caacgccttc	actcctacaa	1320
ccattcctga	tggcatggga	gccaacattc	ccatgatggg	caccacatg	ccaatggctg	1380
gagacatgaa	tggactcagc	cccacccagg	cactcccttc	cccactctcc	atgccatcca	1440
cctcccactg	cacaccccca	cctccgtatc	ccacagattg	cagcattgtc	agtttcttag	1500
cgaggttggg	ctgttcatca	tgtctggact	atttcacgac	ccaggggctg	accaccatct	1560
atcagattga	gcattactcc	atggatgata	tggcaagtct	gaaaatccct	gagcaatttc	1620
gacatgcgat	ctggaagggc	atcctggacc	accggcagct	ccacgaattc	tcctcccctt	1680
ctcatctcct	gcggaaccca	agcagtgcct	ctacagtcag	tgtgggctcc	agtgagaccc	1740
ggggtgagcg	tgttattgat	gctgtgcgat	tcacctcccg	ccagaccatc	tccttcccac	1800
cccagatgta	gtggaatgac	ttcaactttg	acatggatgc	tcgccgcaat	aagcaacagc	1860
gcatcaaaag	ggagggggag	tgagcctcac	catgtgagct	cttcctatcc	ctctcctaac	1920
tgccagcccc	ctaaaagcac	tccctgttaa	tcttcaaagg	cttctcccta	gctcctcccc	1980
ttctcttgtt	ctgatttctt	aggggaagga	gaagtgaagc	gcttacttct	taccctaacc	2040
atctgacctg	gcattctaatt	ctgattctgg	ctttaagcct	tcaaaactat	agcttgacga	2100
actgtagctt	gccatggcta	ggtagaagtg	agcaaaaaag	agttgggtgt	ctccttaage	2160
tgcagagatt	tctcattgac	ttttataaag	catgttcacc	cttatagtct	aagactatat	2220
atataaatgt	ataaatatac	agtatagatt	tttgggtggg	gggcattgag	tattgtttaa	2280
aatgtaattt	aatgaaaga	aaattgagtt	gcacttattg	accatttttt	aattttactg	2340
ttttggatgg	cttgtctata	ctccttcctt	taaggggtat	catgtatggt	gataggtatc	2400
tagagcttaa	tgctacatgt	gagtgacgat	gatgtacaga	ttctttcagt	tcctttggatt	2460
ctaaatacat	gccacatcaa	acctttgagt	agatccattt	ccattgctta	ttatgtaggt	2520
aagactgtag	atatgtattc	ttttctcagt	gttggtatat	tttatattac	tgacatttct	2580
tctagtgatg	atggttcacg	ttgggggtgat	ttaatccagt	tataagaaga	agttcatgtc	2640
caaacgtcct	ctttagtttt	tggttgggaa	tgaggaaaat	tcttaaaaag	cccatagcag	2700
ccagttcaaa	aacacccgac	gtcatgtatt	tgagcatatc	agtaaccccc	ttaaatttaa	2760
taccagatac	cttatcttac	aatattgatt	gggaaaacat	ttgctgccat	tacagaggta	2820
ttaaaactaa	atcttactac	tagattgact	aactcaaata	cacattttgt	actgtttgta	2880
gaattctgat	tgatttgatt	gggatgaatg	ccatctatct	agttctaaca	gtgaagtttt	2940
actgtctatt	aatattcagg	gtaaaatagga	atcattcaga	aatgttgagt	ctgtactaaa	3000
cagtaagata	tctcaatgaa	ccataaattc	aactttgtaa	aaatcttttg	aagcatagat	3060
aatattgttt	ggtaaatgtt	tcttttgttt	ggtaaatgtt	tcytttaaaag	accttcttat	3120

```

tctataaaac tctgcatgta gaggcttggt tacctttctc tctctaaggt ttacaatagg 3180
agtgggtgatt tgaaaaatat aaaattatga gattgggtttt cctgtggcat aaattgcatc 3240
actgtatcat tttctttttt aaccggtaag agtttcagtt tgttggaag taactgtgag 3300
aaccagttt cccgtccatc tcccttaggg actaccata gacatgaaag gtccccacag 3360
agcaagagat aagtctttca tggctgctgt tgcttaaacc acttaaacga agagtccct 3420
tgaaactttg ggaaaacatg ttaatgacaa tattccagat ctttcagaaa tataacacat 3480
ttttttgcat gcatgcaa atgagctctgaa atcttcccat gcattctggt caagggctgt 3540
cattgcacat aagcttccat ttttaatttta aagtgcacaaa gggccagcgt ggctctaaaa 3600
ggtaatgtgt ggattgcctc tgaaaagtgt gtatatattt tgtgtgaaat tgcatacttt 3660
gtattttgat tatttttttt ttcttcttgg gatagtggga tttccagaac cacacttgaa 3720
accttttttt atcgttttttg tattttcatg aaaataccat ttagtaagaa taccacatca 3780
aataagaaat aatgctacaa ttttaagagg ggaggggaagg gaaagttttt ttttttatta 3840
tttttttaaa attttgtatg ttaaagagaa tgagtccttg atttcaaagt tttgtgttac 3900
ttaaagtgtg ataagcactg taaacttctg caacaagcat gcagctttgc aaaccatta 3960
aggggaagaa tgaaagctgt tccctgggcc tagtaagaag acaaaactgct tcccttactt 4020
tgctgagggg ttgaataaac ctaggacttc cgagctatgt cagtactatt caggtaacac 4080
tagggccttg gaaatccctg tactgtgtct catggatttg gcactagcca aagcgaggca 4140
ccccttactg gcttacctcc tcatggcagc ctactctctc tgagtgtatg agtagccagg 4200
gtaaggggta aaaggatagt aagcatagaa accactagaa agtgggctta atggagtctt 4260
tgtggcctca gctcaatgca gttagctgaa gaattgaaaa gtttttgttt ggagacgttt 4320
ataaacagaa atggaaagca gagttttcat taaatccttt tacctttttt ttttcttggt 4380
aatcccctaa aataacagta tgtgggatat tgaatgttaa agggatattt ttttctatta 4440
tttttataat tgtacaaaat taagcaaagt ttaaaagttt tatatgcttt attaatgttt 4500
tcaaaaggta ttatacatgt gatacatttt ttaagcttca gttgcttgct ttctggtact 4560
ttctgttatg ggcttttggg gagccagaag ccaatctaca atctcttttt gtttgccagg 4620
acatgcaata aaatttaaaa aataaataaa aacta 4655

```

```

<210> 152
<211> 586
<212> PRT
<213> Homo sapien

```

```

<400> 152

```

```

Met Leu Tyr Leu Glu Asn Asn Ala Gln Thr Gln Phe Ser Glu Pro Gln
1          5          10          15
Tyr Thr Asn Leu Gly Leu Leu Asn Ser Met Asp Gln Gln Ile Gln Asn
20          25          30
Gly Ser Ser Ser Thr Ser Pro Tyr Asn Thr Asp His Ala Gln Asn Ser
35          40          45
Val Thr Ala Pro Ser Pro Tyr Ala Gln Pro Ser Ser Thr Phe Asp Ala
50          55          60
Leu Ser Pro Ser Pro Ala Ile Pro Ser Asn Thr Asp Tyr Pro Gly Pro
65          70          75          80
His Ser Phe Asp Val Ser Phe Gln Gln Ser Ser Thr Ala Lys Ser Ala
85          90          95
Thr Trp Thr Tyr Ser Thr Glu Leu Lys Lys Leu Tyr Cys Gln Ile Ala
100         105         110
Lys Thr Cys Pro Ile Gln Ile Lys Val Met Thr Pro Pro Pro Gln Gly
115         120         125
Ala Val Ile Arg Ala Met Pro Val Tyr Lys Lys Ala Glu His Val Thr
130         135         140
Glu Val Val Lys Arg Cys Pro Asn His Glu Leu Ser Arg Glu Phe Asn
145         150         155         160
Glu Gly Gln Ile Ala Pro Ser Ser His Leu Ile Arg Val Glu Gly Asn

```

					165					170					175
Ser	His	Ala	Gln	Tyr	Val	Glu	Asp	Pro	Ile	Thr	Gly	Arg	Gln	Ser	Val
			180					185					190		
Leu	Val	Pro	Tyr	Glu	Pro	Pro	Gln	Val	Gly	Thr	Glu	Phe	Thr	Thr	Val
		195					200					205			
Leu	Tyr	Asn	Phe	Met	Cys	Asn	Ser	Ser	Cys	Val	Gly	Gly	Met	Asn	Arg
	210					215					220				
Arg	Pro	Ile	Leu	Ile	Ile	Val	Thr	Leu	Glu	Thr	Arg	Asp	Gly	Gln	Val
225					230					235					240
Leu	Gly	Arg	Arg	Cys	Phe	Glu	Ala	Arg	Ile	Cys	Ala	Cys	Pro	Gly	Arg
				245					250					255	
Asp	Arg	Lys	Ala	Asp	Glu	Asp	Ser	Ile	Arg	Lys	Gln	Gln	Val	Ser	Asp
			260					265					270		
Ser	Thr	Lys	Asn	Gly	Asp	Gly	Thr	Lys	Arg	Pro	Phe	Arg	Gln	Asn	Thr
		275					280					285			
His	Gly	Ile	Gln	Met	Thr	Ser	Ile	Lys	Lys	Arg	Arg	Ser	Pro	Asp	Asp
	290					295					300				
Glu	Leu	Val	Tyr	Leu	Pro	Val	Arg	Gly	Arg	Glu	Thr	Tyr	Glu	Met	Leu
305					310					315					320
Val	Lys	Ile	Lys	Glu	Ser	Leu	Glu	Leu	Met	Gln	Tyr	Leu	Leu	Gln	His
				325					330					335	
Thr	Ile	Glu	Thr	Tyr	Arg	Gln	Gln	Gln	Gln	Gln	Gln	His	Gln	His	Leu
			340					345					350		
Leu	Gln	Lys	Gln	Thr	Ser	Ile	Gln	Ser	Pro	Ser	Ser	Tyr	Gly	Asn	Ser
		355					360					365			
Ser	Pro	Pro	Leu	Asn	Lys	Met	Asn	Ser	Met	Asn	Lys	Leu	Pro	Ser	Val
	370					375					380				
Ser	Gln	Leu	Ile	Asn	Pro	Gln	Gln	Arg	Asn	Ala	Leu	Thr	Pro	Thr	Thr
385				390						395					400
Ile	Pro	Asp	Gly	Met	Gly	Ala	Asn	Ile	Pro	Met	Met	Gly	Thr	His	Met
			405						410					415	
Pro	Met	Ala	Gly	Asp	Met	Asn	Gly	Leu	Ser	Pro	Thr	Gln	Ala	Leu	Pro
			420					425					430		
Pro	Pro	Leu	Ser	Met	Pro	Ser	Thr	Ser	His	Cys	Thr	Pro	Pro	Pro	Pro
		435					440					445			
Tyr	Pro	Thr	Asp	Cys	Ser	Ile	Val	Ser	Phe	Leu	Ala	Arg	Leu	Gly	Cys
	450					455					460				
Ser	Ser	Cys	Leu	Asp	Tyr	Phe	Thr	Thr	Gln	Gly	Leu	Thr	Thr	Ile	Tyr
465				470						475					480
Gln	Ile	Glu	His	Tyr	Ser	Met	Asp	Asp	Leu	Ala	Ser	Leu	Lys	Ile	Pro
				485					490					495	
Glu	Gln	Phe	Arg	His	Ala	Ile	Trp	Lys	Gly	Ile	Leu	Asp	His	Arg	Gln
			500					505					510		
Leu	His	Glu	Phe	Ser	Ser	Pro	Ser	His	Leu	Leu	Arg	Thr	Pro	Ser	Ser
		515					520								

<210> 153
 <211> 2007
 <212> DNA
 <213> Homo sapien

<400> 153

gaattcgtcg	ctgctccagg	gaaagttctg	ttactccact	gactctctct	tttctgata	60
acatggccag	caagaaagta	attacagtgt	ttggagcaac	aggagctcaa	ggtggctctg	120
tggccagggc	aattttggag	agcaaaaaat	ttgcagtggag	agcagtgacc	agggatgtga	180
cttgaccaa	tgccctggag	ctccagcgcc	ttggagctga	ggtggtcaaa	ggtgacctga	240
atgataaagc	atcggtggac	agtgccttaa	aaggtgtcta	tggggccttc	ttggtgacca	300
acttctggga	ccctctcaac	caagataagg	aagtgtgtcg	ggggaagctg	gtggcagact	360
ccgccaagca	cctgggtctg	aagcacgtgg	tgtacagcgg	cctggagaac	gtcaagcgac	420
tgacggatgg	caagctggag	gtgccgcact	ttgacagcaa	gggcgagggtg	gaggagtact	480
tctggtccat	tggcatcccc	atgaccagtg	tccgcgtggc	ggcctacttt	gaaaactttc	540
tgcggcgctg	gcggcccggtg	aaagcctctg	atggagatta	ctacaccttg	gctgtaccga	600
tgggagatgt	accaatggat	ggtatctctg	ttgctgatat	tggagcagcc	gtctctagca	660
tttttaattc	tccagaggaa	tttttaggca	aggcctgggg	gtcagtgcga	gaagcactaa	720
caatacagca	atatgctgat	gttttgtcca	aggctttggg	gaaagaagtc	cgagatgcaa	780
agattacccc	ggaagctttc	gagaagctgg	gattccctgc	agcaaaggaa	atagccaata	840
tgtgtcgttt	ctatgaaatg	aagccagacc	gagatgtcaa	tctcaccac	caactaaatc	900
ccaaagtcaa	aagcttcagc	cagtttatct	cagagaacca	gggagccttc	aagggcattgt	960
agaaaatcag	ctgttcagat	aggcctctgc	accacacagc	ctctttcctc	tctgatcctt	1020
ttcctcttta	cggcacaaca	ttcatgttga	cagaacatgc	tggaatgcaa	ttgtttgcaa	1080
caccgaagga	tttctgctgg	tgcctctctc	agtaggaagc	actgcattgg	tgataggaca	1140
cggtaatttg	attcacattt	aacttgctag	ttagtataaa	gggtggtaca	actgtttggt	1200
aaaatgagaa	gcctcggaac	ttggagcttc	tctcctacca	ctaattgggag	ggcagattat	1260
actgggattt	ctcctgggtg	agtaatttca	agccctaattg	ctgaaattcc	cctaggcagc	1320
tccagttttc	tcaactgcat	tgcaaaaattc	ccagtgaact	tttaagtact	tttaacttaa	1380
aaaaatgaac	atctttgtag	agaattttct	ggggaacatg	gtgttcaatg	aacaagcaca	1440
agcattggaa	atgctaaaaat	tcagttttgc	ctcaagattg	gaagtttatt	ttctgactca	1500
ttcatgaagt	catctattga	gccaccattc	aattattcat	ctattaattc	cttgatcctt	1560
catttatcca	ttctgcaaac	ttttcttgag	caccagcacg	ggtggccatt	tgtggacttc	1620
tcttcattcc	tatgtgtttt	cttatcaaa	tgatccactc	tcgaaaaggct	cotttccagt	1680
ctgtggttgg	gttcaagtca	tgccaggggc	agggggccca	tctcctcggt	tagctctagg	1740
caaaaaccag	gggatctgca	ctggggagcg	ggggcaggaa	gctggaggga	aggcctgtga	1800
agggtaggga	tgtggaaaga	caaggtgaca	gaaggaccca	ataggacctt	tctatatctc	1860
tggcttagca	ttttctacat	catattgtaa	tcgtcttatt	tgctagtttt	cttcttact	1920
gtgagtgact	aacagtcctc	tttatccag	tgcttggtac	ataataagtg	atcaataaat	1980
gttgattgac	taaaaaaaaa	aaaaaaa				2007

<210> 154
 <211> 2148
 <212> DNA
 <213> Homo sapien

<400> 154

gaattcgtcg	ctgctccagg	gaaagttctg	ttactccact	gactctctct	tttctgata	60
acatggccag	caagaaagta	attacagtgt	ttggagcaac	aggagctcaa	ggtggctctg	120
tggccagggc	aattttggag	agcaaaaaat	ttgcagtggag	agcagtgacc	agggatgtga	180
cttgaccaa	tgccctggag	ctccagcgcc	ttggagctga	ggtggtcaaa	ggtgacctga	240
atgataaagc	atcggtggac	agtgccttaa	aaggggaagc	tgggtggcaga	ctccgccaag	300
cacctgggtc	tgaagcagct	ggtgtacagc	ggcctggaga	acgtcaagcg	actgacggat	360
ggcaagctgg	aggtgccgca	ctttgacagc	aagggcgagg	tggaggagta	cttctggtcc	420

```

attggcatcc ccatgaccag tgtccgcgtg ggggcctact ttgaaaactt tctcgcggcg 480
tggcgggccc tgaaagcctc tgatggagat tactacacct tggctgtacc gatgggagat 540
gtaccaatgg atggtatctc tgttgctgat attggagcag ccgtctctag cttttttaat 600
tctccagagg aatttttagg caaggccgtg gggctcagtg cagaagcact aacaatacag 660
caatatgctg atgttttgtc caaggctttg gggaaagaag tccgagatgc aaagactatc 720
tgtgctatag atgaccagaa aacagtggaa gaaggtttca tgggaagacgt gggcttgagt 780
tggtccttga gggaacatga ccatgtatag acagaggagg catcaagaag gctggcctgg 840
ctaattctgg aataaacacg acaaaccaga ggcagtacgg gaaggaggca aattctggct 900
ctgcctctat ccttgattac cccggaagct ttcgagaagc tgggattccc tgcagcaaag 960
gaaatagcca atatgtgtcg tttctatgaa atgaagccag accgagatgt caatctcacc 1020
caccaactaa atcccaaagt caaaagcttc agccatttta tctcagagaa ccaggaggcc 1080
ttcaagggca tgtagaaaat cagctgttca gataggcctc tgcaccacac agcctctttc 1140
ctctctgata cttttcctct ttacggcaca acattcatgt tgacagaaca tgctggaatg 1200
caattgtttg caacaccgaa ggatttcctg cggctgcctc ttcagtagga agcactgcat 1260
tggatgatagg acacggtaat ttgattcaca ttttaacttgc tagttagtga taagggtggg 1320
acaactgttt ggtaaaatga gaagcctcgg aacttgagc ttctctccta ccactaatgg 1380
gagggcagat tatactggga tttctcctgg gtgagtaatt tcaagcccta atgctgaaat 1440
tcccctaggc agctccagtt ttctcaactg cattgcaaaa ttcccagtg aacttttaagt 1500
acttttaact taaaaaaatg aacatctttg tagagaattt tctggggaac atgggtgttca 1560
atgaacaagc acaagcattg gaaatgctaa aattcagttt tgcctcaaga ttggaagttt 1620
attttctgac tcattcatga agtcatctat tgagccacca ttcaattatt catctattaa 1680
ttccttgata cttcatttat ccattctgca aacttttctt gagcaccagc acgggtggcc 1740
atttgtggac ttctcttcat tcctatgtgt tttcttatca aagtgatcca ctctcgaaag 1800
gctcctttcc agtctgtggt tgggttcaag tcatgccagg gccagggggc ccactcctc 1860
gtttagctct aggcaaaatc caggggatct gcagtgggga gcgggggcag gaagctggag 1920
ggaaggcctg tgaagggtag ggatgtggaa agacaagggt acagaaggac ccaataggac 1980
ctttctatat ctctggctta gcattttcta catcatattg taatcgtctt atttgctagt 2040
tttcttcctt actgtgagtg actaacagtc atctttatcc cagtgcctgg tacataataa 2100
gtgatcaata aatgttgatt gactaaatga aaaaaaaaaa aaaaaaaa 2148

```

<210> 155

<211> 153

<212> PRT

<213> Homo sapien

<400> 155

```

Met Thr Ser Val Arg Val Ala Ala Tyr Phe Glu Asn Phe Leu Ala Ala
1          5          10          15
Trp Arg Pro Val Lys Ala Ser Asp Gly Asp Tyr Tyr Thr Leu Ala Val
20        25        30
Pro Met Gly Asp Val Pro Met Asp Gly Ile Ser Val Ala Asp Ile Gly
35        40        45
Ala Ala Val Ser Ser Ile Phe Asn Ser Pro Glu Glu Phe Leu Gly Lys
50        55        60
Ala Val Gly Leu Ser Ala Glu Ala Leu Thr Ile Gln Gln Tyr Ala Asp
65        70        75        80
Val Leu Ser Lys Ala Leu Gly Lys Glu Val Arg Asp Ala Lys Ile Thr
85        90        95
Pro Glu Ala Phe Glu Lys Leu Gly Phe Pro Ala Ala Lys Glu Ile Ala
100       105       110
Asn Met Cys Arg Phe Tyr Glu Met Lys Pro Asp Arg Asp Val Asn Leu
115       120       125
Thr His Gln Leu Asn Pro Lys Val Lys Ser Phe Ser Gln Phe Ile Ser
130       135       140

```

Glu Asn Gln Gly Ala Phe Lys Gly Met
145 150

<210> 156
<211> 128
<212> PRT
<213> Homo sapien

<400> 156
Met Thr Ser Val Arg Val Ala Ala Tyr Phe Glu Asn Phe Leu Ala Ala
1 5 10 15
Trp Arg Pro Val Lys Ala Ser Asp Gly Asp Tyr Tyr Thr Leu Ala Val
20 25 30
Pro Met Gly Asp Val Pro Met Asp Gly Ile Ser Val Ala Asp Ile Gly
35 40 45
Ala Ala Val Ser Ser Ile Phe Asn Ser Pro Glu Glu Phe Leu Gly Lys
50 55 60
Ala Val Gly Leu Ser Ala Glu Ala Leu Thr Ile Gln Gln Tyr Ala Asp
65 70 75 80
Val Leu Ser Lys Ala Leu Gly Lys Glu Val Arg Asp Ala Lys Thr Ile
85 90 95
Cys Ala Ile Asp Asp Gln Lys Thr Val Glu Glu Gly Phe Met Glu Asp
100 105 110
Val Gly Leu Ser Trp Ser Leu Arg Glu His Asp His Val Ala Gly Ala
115 120 125

<210> 157
<211> 424
<212> DNA
<213> Homo sapien

<220>
<221> misc_feature
<222> (1)...(424)
<223> n = A,T,C or G

<400> 157
ctgcagcccg ggggatccac tagtccagtg tgggtggaatt catttgtctt tacaagactt 60
ggatacatta cagcagacat ggaaatataa ttttaaaaaa tttctctcca acctccttca 120
aattcagtca ccaactgttat attaccttct ccaggaacco tccagtgggg aaggctgcga 180
tattagattt ccttgtatgc aaagtttttg ttgaaagctg tgctcagagg aggtgagagg 240
agaggaagga gaaaactgca tcataacttt acagaattga atctagagtc ttccccgaaa 300
agcccagaaa cttctctgcn gnatctggct tgtccatctg gtctaagggt gctgcttctt 360
ccccagccat cgagtcagtt tgtgcccacg aataatacac gacctgctat ttcccatgac 420
tgct 424

<210> 158
<211> 2099
<212> DNA
<213> Homo sapien

<400> 158
ccgcggttaa aaggcgagc aggtgggagc cggggccttc acccgaaacc cgacgagagc 60
ccgacagccg gggcgcccg agcccagcct gcctgccag ccggagcgaa gggcgccgc 120

```

ccgcgcagag cccgcgccag gccgcgccgc cgcagagcag ttaaaacgtg caggcaccag 180
aaggcacttc ctgtcgggtga agaagacctg tctccgggtg cacgggcata ctgtgttttg 240
caaacggggc tgacctccct tcctggggag caggaagggt caggggaagga aaagaagtac 300
agaagatctg gctaaacaat ttctgtatgg cgaaagaaaa attctaactt gtacgccctc 360
ttcatgcata tttaattcaa ttggaatatt ccaggcgaca tctcactga ccgagcaaag 420
attgacattc gtatcatcac tgtgcaccat tggcttctag gcactccagt ggggtaggag 480
aaggaggtct gaaaccctcg cagagggatc ttgccctcat tctttgggtc tgaaacactg 540
gcagtcgttg gaaacaggac tcagggataa accagcgcaa tggattgggg gacgctgcac 600
actttcatcg ggggtgtcaa caaacactcc accagcatcg ggaagggtgt gatcacagtc 660
atctttatct tccgagtcac gatcctcggt gtggctgccc aggaagtgtg ggggtgacgag 720
caagaggact tcgtctgcaa cacactgcaa ccgggatgca aaaatgtgtg ctatgaccac 780
tttttcccg tgtcccatc ccggctgtgg gccctccagc tgatcttcgt ctccaccca 840
gcgtctgtgg tggccatgca tgtggcctac tacaggcacg aaaccactcg caagttcagg 900
cgaggagaga agaggaatga tttcaaagac atagaggaca ttaaaaagca gaaggttcgg 960
atagaggggt cgctgtgggt gacgtacacc agcagcatct ttttcgaat catctttgaa 1020
gcagccttta tgtatgtgtt ttacttcctt tacaatgggt accacctgcc ctgggtgttg 1080
aaatgtggga ttgacctcg ccccaacctt gttgactgct ttatttctag gccaacagag 1140
aagaccgtgt ttaccatttt tatgatttct gcgtctgtga tttgcatgct gcttaacgtg 1200
gcagagttgt gctacctgct gctgaaagtg tgttttagga gatcaaagag agcacagacg 1260
caaaaaaatc accccaatca tgcctaagag gagagtaagc agaatgaaat gaatgagctg 1320
atttcagata gtggtcaaaa tgcaatcaca ggttcccaag ctaaactttt caaggtaaaa 1380
tgtagctgcg tcataaggag acttctgtct tctccagaag gcaataccaa cctgaaagt 1440
ccttctgtag cctgaagagt ttgtaaatga ctttcataat aaatagacac ttgagttaac 1500
tttttgtagg atacttgctc cattcataca caacgtaatc aaatatgtgg tccatctctg 1560
aaaacaagag actgcttgac aaaggagcat tgcagtcact ttgacagggt ccttttaagt 1620
ggactctctg acaaagtggg tactttctga aaatttatat aactgttggt gataaggaac 1680
atztatccag gaattgatac gtttattagg aaaagatat tttataggct tggatgtttt 1740
tagttctgac tttgaattta tataaagtat ttttataatg actgggtctt cttacctgga 1800
aaaacatgcg atgttagttt tagaattaca ccacaagtat ctaaatttgg aacttacaaa 1860
gggtctatct tgtaaattat gttttgcatt gtctgttggc aaatttgtga actgtcatga 1920
tacgcttaag gtggaagtgt ttcattgcac aatatatctt tactgctttc tgaatgtaga 1980
cggaacagtg tggaagcaga aggttttttt aactcatccg tttgccaatc attgcaaaca 2040
actgaaatgt ggatgtgatt gcctcaataa agctcgtccc cattgcttaa aaaaaaaaaa 2099

```

<210> 159

<211> 291

<212> PRT

<213> Homo sapien

<400> 159

```

Met Asp Trp Gly Thr Leu His Thr Phe Ile Gly Gly Val Asn Lys His
1          5          10          15
Ser Thr Ser Ile Gly Lys Val Trp Ile Thr Val Ile Phe Ile Phe Arg
20          25          30
Val Met Ile Leu Val Val Ala Ala Gln Glu Val Trp Gly Asp Glu Gln
35          40          45
Glu Asp Phe Val Cys Asn Thr Leu Gln Pro Gly Cys Lys Asn Val Cys
50          55          60
Tyr Asp His Phe Phe Pro Val Ser His Ile Arg Leu Trp Ala Leu Gln
65          70          75          80
Leu Ile Phe Val Ser Thr Pro Ala Leu Leu Val Ala Met His Val Ala
85          90          95
Tyr Tyr Arg His Glu Thr Thr Arg Lys Phe Arg Arg Gly Glu Lys Arg
100          105          110

```

Asn Asp Phe Lys Asp Ile Glu Asp Ile Lys Lys Gln Lys Val Arg Ile
 115 120 125
 Glu Gly Ser Leu Trp Trp Thr Tyr Thr Ser Ser Ile Phe Phe Arg Ile
 130 135 140
 Ile Phe Glu Ala Ala Phe Met Tyr Val Phe Tyr Phe Leu Tyr Asn Gly
 145 150 155 160
 Tyr His Leu Pro Trp Val Leu Lys Cys Gly Ile Asp Pro Cys Pro Asn
 165 170 175
 Leu Val Asp Cys Phe Ile Ser Arg Pro Thr Glu Lys Thr Val Phe Thr
 180 185 190
 Ile Phe Met Ile Ser Ala Ser Val Ile Cys Met Leu Leu Asn Val Ala
 195 200 205
 Glu Leu Cys Tyr Leu Leu Leu Lys Val Cys Phe Arg Arg Ser Lys Arg
 210 215 220
 Ala Gln Thr Gln Lys Asn His Pro Asn His Ala Leu Lys Glu Ser Lys
 225 230 235 240
 Gln Asn Glu Met Asn Glu Leu Ile Ser Asp Ser Gly Gln Asn Ala Ile
 245 250 255
 Thr Gly Ser Gln Ala Lys His Phe Lys Val Lys Cys Ser Cys Val Ile
 260 265 270
 Arg Arg Leu Leu Ser Ser Pro Glu Gly Asn Thr Asn Leu Lys Val Pro
 275 280 285
 Ser Val Ala
 290

<210> 160

<211> 3951

<212> DNA

<213> Homo sapien

<400> 160

tctgcatcca	tattgaaaac	ctgacacaat	gtatgcagca	ggctcagtgt	gagtgaactg	60
gaggcttctc	tacaacatga	cccaaaggag	cattgcaggt	cctatttgca	acctgaagtt	120
tgtgactctc	ctgggttgct	taagttcaga	actcccattc	ctgggagctg	gagtacagct	180
tcaagacaat	gggtataatg	gattgctcat	tgcaattaat	cctcaggtac	ctgagaatca	240
gaacctcatc	tcaaacatta	aggaaatgat	aactgaagct	tcattttacc	tattttaatgc	300
taccaagaga	agagtatttt	tcagaaatat	aaagatttta	atacctgcca	catggaaagc	360
taataataac	agcaaaataa	aacaagaatc	atatgaaaag	gcaaatgtca	tagtgactga	420
ctggtatggg	gcacatggag	atgatccata	caccctacaa	tacagagggt	gtggaaaaga	480
gggaaaatac	attcatttca	cacctaattt	cctactgaat	gataacttaa	cagctggcta	540
cggatcacga	ggccgagtgt	ttgtccatga	atgggcccac	ctccgttggg	gtgtgttcga	600
tgagtataac	aatgacaaac	ctttctacat	aaatgggcaa	aatcaaatta	aagtgacaag	660
gtgttcatct	gacatcacag	gcatttttgt	gtgtgaaaaa	ggtccttgcc	ccaagaaaaa	720
ctgtattatt	agtaagcttt	ttaaagaagg	atgcaccttt	atctacaata	gcacccaaaa	780
tgcaactgca	tcaataatgt	tcatgcaaag	tttatcttct	gtggttgaat	tttgtaatgc	840
aagtacccac	aaccaagaag	caccaaacct	acagaaccag	atgtgcagcc	tcagaagtgc	900
atgggatgta	atcacagact	ctgctgactt	tcaccacagc	tttcccatga	acgggactga	960
gcttccacct	cctcccacat	tctcgcttgt	agaggctggg	gacaaaagtg	tctgtttagt	1020
gctggatgtg	tccagcaaga	tggcagaggc	tgacagactc	cttcaactac	aacaagccgc	1080
agaattttat	ttgatgcaga	ttgttgaaat	tcataccttc	gtgggcattg	ccagtttcga	1140
cagcaaagga	gagatcagag	cccagctaca	ccaaattaac	agcaatgatg	atcgaaagtt	1200
gctggtttca	tatctgccca	ccactgtatc	agctaaaaca	gacatcagca	tttgttcagg	1260
gcttaagaaa	ggatttgagg	tggttgaaaa	actgaatgga	aaagcttatg	gctctgtgat	1320
gatattagtg	accagcggag	atgataagct	tcttggcaat	tgcttaccga	ctgtgctcag	1380


```

cagtgggttca acaattcact ccattgccct gggttcatct gcagccccaa atctggagga 1440
attatcacgt cttacaggag gtttaaagtt ctttggtcca gatatatcaa actccaatag 1500
catgattgat gctttcagta gaatttcctc tggaaactgga gacattttcc agcaacatat 1560
tcagcttgaa agtacaggtg aaaatgtcaa acctcaccat caattgaaaa acacagtgaac 1620
tgtggataat actgtgggca acgacactat gtttctagtt acgtggcagg ccagtgggtcc 1680
tcctgagatt atattatttg atcctgatgg acgaaaatac tacacaaata attttatcac 1740
caatctaact tttcggacag ctagtctttg gattccagga acagctaagc ctgggcactg 1800
gacttacacc ctgaacaata cccatcattc tctgcaagcc ctgaaagtga cagtgcctc 1860
tcgcgcctcc aactcagctg tgccccccagc cactgtggaa gcctttgtgg aaagagacag 1920
cctccatttt cctcatcctg tgatgattta tgccaatgtg aaacagggat tttatcccat 1980
tcttaatgcc actgtcactg ccacagttga gccagagact ggagatcctg ttacgtgag 2040
actccttgat gatggagcag gtgctgatgt tataaaaaat gatggaattt actcgaggta 2100
ttttttctcc tttgctgcaa atggtagata tagcttgaaa gtgcatgtca atcactctcc 2160
cagcataagc accccagccc actctattec agggagtcac gctatgtatg taccaggtta 2220
cacagcaaac ggtaatatcc agatgaatgc tccaaggaaa tcagtaggca gaaatgagga 2280
ggagcgaaaag tggggcttta gccgagtcag ctccaggaggc tccttttcag tgctgggagt 2340
tcagctggc cccaccctg atgtgtttcc accatgcaaa attattgacc tggaagctgt 2400
aaaagtagaa gaggaattga ccctatcttg gacagcacct ggagaagact ttgatcaggg 2460
ccaggctaca agctatgaaa taagaatgag taaaagtcta cagaatatcc aagatgactt 2520
taacaatgct attttagtaa atacatcaaa gcgaaatcct cagcaagctg gcatcaggga 2580
gatattttacg ttctcaccct aaatttccac gaatggacct gaacatcagc caaatggaga 2640
aacacatgaa agccacagaa tttatgttgc aatacgagca atggatagga actccttaca 2700
gtctgtgtga tctaacattg cccaggcgcc tctgtttatt cccccaatt ctgatcctgt 2760
acctgccaga gattatctta tattgaaagg agttttaaca gcaatgggtt tgataggaat 2820
catttgcctt attatagttg tgacacatca tactttaagc aggaaaaaga gagcagacaa 2880
gaaagagaat ggaacaaaat tattataaat aaatatccaa agtgtcttcc ttcttagata 2940
taagacccat ggccttcgac taaaaaaaca tactaacaaa gtcaaattaa catcaaaact 3000
gtattaaaaat gcattgagtt tttgtacaat acagataaga tttttacatg gtagatcaac 3060
aaattctttt tgggggtaga ttagaaaacc cttacacttt ggctatgaac aaataataaa 3120
aattattctt taaagtaatg tctttaaaagg caaagggaag ggtaaagtgc gaccagtgtc 3180
aaggaaagt tgttttattg aggtggaaaa atagccccaa gcagagaaaa ggagggtagg 3240
tctgcattat aactgtctgt gtgaagcaat catttagtta ctttgattaa tttttctttt 3300
ctccttatct gtgcagaaca ggttgcttgt ttacaactga agatcatgct atatttcata 3360
tatgaagccc ctaatgcaaa gctctttacc tcttgctatt ttgttatata tattacagat 3420
gaaatctcac tgctaatgct cagagatctt ttttactgt aagaggtaac ctttaacaat 3480
atgggtatta cctttgtctc ttcaaccgg ttttatgaca aaggtctatt gaattttatt 3540
gtttgtaagt ttctactccc atcaaagcag ctttttaagt tattgccttg gttattatgg 3600
atgatagtta tagcccttat aatgccttaa ctaagggaaga aaagatgtta ttctgagttt 3660
gttttaatac atatatgaac atatagtttt attcaattaa accaaagaag aggtcagcag 3720
ggagatacta acctttggaa atgattagct ggctctgttt tttggttaaa taagagtctt 3780
taatcctttc tccatcaaga gttacttacc aagggcaggg gaagggggat atagagggtcc 3840
caaggaaata aaaatcatct ttcatcttta attttactcc ttctcttat ttttttaaaa 3900
gattatcgaa caataaaatc atttgccttt ttaattaaaa acataaaaaa a 3951

```

<210> 161
 <211> 943
 <212> PRT
 <213> Homo sapien

<400> 161

```

Met Thr Gln Arg Ser Ile Ala Gly Pro Ile Cys Asn Leu Lys Phe Val
 1           5           10           15
Thr Leu Leu Val Ala Leu Ser Ser Glu Leu Pro Phe Leu Gly Ala Gly
 20           25           30

```

Val	Gln	Leu	Gln	Asp	Asn	Gly	Tyr	Asn	Gly	Leu	Leu	Ile	Ala	Ile	Asn	
	35						40					45				
Pro	Gln	Val	Pro	Glu	Asn	Gln	Asn	Leu	Ile	Ser	Asn	Ile	Lys	Glu	Met	
	50					55					60					
Ile	Thr	Glu	Ala	Ser	Phe	Tyr	Leu	Phe	Asn	Ala	Thr	Lys	Arg	Arg	Val	
65					70					75					80	
Phe	Phe	Arg	Asn	Ile	Lys	Ile	Leu	Ile	Pro	Ala	Thr	Trp	Lys	Ala	Asn	
			85						90					95		
Asn	Asn	Ser	Lys	Ile	Lys	Gln	Glu	Ser	Tyr	Glu	Lys	Ala	Asn	Val	Ile	
			100					105					110			
Val	Thr	Asp	Trp	Tyr	Gly	Ala	His	Gly	Asp	Asp	Pro	Tyr	Thr	Leu	Gln	
	115						120					125				
Tyr	Arg	Gly	Cys	Gly	Lys	Glu	Gly	Lys	Tyr	Ile	His	Phe	Thr	Pro	Asn	
	130					135					140					
Phe	Leu	Leu	Asn	Asp	Asn	Leu	Thr	Ala	Gly	Tyr	Gly	Ser	Arg	Gly	Arg	
145					150					155					160	
Val	Phe	Val	His	Glu	Trp	Ala	His	Leu	Arg	Trp	Gly	Val	Phe	Asp	Glu	
				165					170					175		
Tyr	Asn	Asn	Asp	Lys	Pro	Phe	Tyr	Ile	Asn	Gly	Gln	Asn	Gln	Ile	Lys	
			180					185					190			
Val	Thr	Arg	Cys	Ser	Ser	Asp	Ile	Thr	Gly	Ile	Phe	Val	Cys	Glu	Lys	
	195						200					205				
Gly	Pro	Cys	Pro	Gln	Glu	Asn	Cys	Ile	Ile	Ser	Lys	Leu	Phe	Lys	Glu	
	210					215					220					
Gly	Cys	Thr	Phe	Ile	Tyr	Asn	Ser	Thr	Gln	Asn	Ala	Thr	Ala	Ser	Ile	
225					230					235					240	
Met	Phe	Met	Gln	Ser	Leu	Ser	Ser	Val	Val	Glu	Phe	Cys	Asn	Ala	Ser	
				245					250					255		
Thr	His	Asn	Gln	Glu	Ala	Pro	Asn	Leu	Gln	Asn	Gln	Met	Cys	Ser	Leu	
		260					265						270			
Arg	Ser	Ala	Trp	Asp	Val	Ile	Thr	Asp	Ser	Ala	Asp	Phe	His	His	Ser	
	275						280					285				
Phe	Pro	Met	Asn	Gly	Thr	Glu	Leu	Pro	Pro	Pro	Pro	Thr	Phe	Ser	Leu	
	290					295					300					
Val	Glu	Ala	Gly	Asp	Lys	Val	Val	Cys	Leu	Val	Leu	Asp	Val	Ser	Ser	
305					310					315					320	
Lys	Met	Ala	Glu	Ala	Asp	Arg	Leu	Leu	Gln	Leu	Gln	Gln	Ala	Ala	Glu	
				325					330					335		
Phe	Tyr	Leu	Met	Gln	Ile	Val	Glu	Ile	His	Thr	Phe	Val	Gly	Ile	Ala	
		340					345						350			
Ser	Phe	Asp	Ser	Lys	Gly	Glu	Ile	Arg	Ala	Gln	Leu	His	Gln	Ile	Asn	
	355					360						365				
Ser	Asn	Asp	Asp	Arg	Lys	Leu	Leu	Val	Ser	Tyr	Leu	Pro	Thr	Thr	Val	
	370					375					380					
Ser	Ala	Lys	Thr	Asp	Ile	Ser	Ile	Cys	Ser	Gly	Leu	Lys	Lys	Gly	Phe	
385					390					395					400	
Glu	Val	Val	Glu	Lys	Leu	Asn	Gly	Lys	Ala	Tyr	Gly	Ser	Val	Met	Ile	
				405					410					415		
Leu	Val	Thr	Ser	Gly	Asp	Asp	Lys	Leu	Gly	Asn	Cys	Leu	Pro	Thr		
		420					425					430				
Val	Leu	Ser	Gly	Ser	Thr	Ile	His	Ser	Ile	Ala	Leu	Gly	Ser	Ser		
	435					440					445					
Ala	Ala	Pro	Asn	Leu	Glu	Glu	Leu	Ser	Arg	Leu	Thr	Gly	Gly	Leu	Lys	
	450					455					460					

Phe 465	Phe 465	Val	Pro	Asp	Ile 470	Ser	Asn	Ser	Asn	Ser	Met	Ile	Asp	Ala	Phe 480
Ser	Arg	Ile	Ser	Ser 485	Gly	Thr	Gly	Asp	Ile 490	Phe	Gln	Gln	His	Ile 495	Gln
Leu	Glu	Ser	Thr	Gly 500	Glu	Asn	Val	Lys	Pro 505	His	His	Gln	Leu	Lys 510	Asn
Thr	Val	Thr	Val	Asp 515	Asn	Thr	Val	Gly 520	Asn	Asp	Thr	Met	Phe	Leu 525	Val
Thr	Trp	Gln	Ala	Ser 530	Gly	Pro	Pro	Glu 535	Ile	Ile	Leu	Phe	Asp	Pro 540	Asp
Gly 545	Arg	Lys	Tyr	Tyr 550	Thr	Asn	Asn	Phe 555	Ile	Thr	Asn	Leu	Thr	Phe 560	Arg
Thr	Ala	Ser	Leu	Trp 565	Ile	Pro	Gly	Thr 570	Ala	Lys	Pro	Gly	His	Trp 575	Thr
Tyr	Thr	Leu	Asn 580	Thr	His	His	Ser 585	Leu	Gln	Ala	Leu	Lys 590	Val	Thr	Thr
Val	Thr	Ser 595	Arg	Ala	Ser	Asn	Ser 600	Val	Pro	Pro	Ala 605	Thr	Val	Glu	Glu
Ala	Phe 610	Val	Glu	Arg	Asp	Ser 615	Leu	His	Phe	Pro	His 620	Pro	Val	Met	Ile
Tyr 625	Ala	Asn	Val	Lys	Gln 630	Gly	Phe	Tyr	Pro	Ile	Leu 635	Asn	Ala	Thr 640	Val
Thr	Ala	Thr	Val	Glu 645	Pro	Glu	Thr	Gly	Asp 650	Pro	Val	Thr	Leu	Arg 655	Leu
Leu	Asp	Asp	Gly 660	Ala	Gly	Ala	Asp 665	Val	Ile	Lys	Asn	Asp 670	Gly	Ile 675	Tyr
Ser	Arg	Tyr 675	Phe	Phe	Ser	Phe	Ala 680	Ala	Asn	Gly	Arg 685	Tyr	Ser	Leu 690	Lys
Val	His 690	Val	Asn	His	Ser	Pro 695	Ser	Ile	Ser	Thr	Pro 700	Ala	His	Ser 705	Ile
Pro	Gly	Ser	His	Ala 710	Met	Tyr	Val	Pro	Gly	Tyr 715	Thr	Ala	Asn	Gly 720	Asn
Ile	Gln	Met	Asn 725	Ala	Pro	Arg	Lys	Ser	Val 730	Gly	Arg	Asn	Glu	Glu 735	Glu
Arg	Lys	Trp	Gly 740	Phe	Ser	Arg	Val	Ser 745	Ser	Gly	Gly	Ser 750	Phe	Ser 755	Val
Leu	Gly	Val 755	Pro	Ala	Gly	Pro	His 760	Pro	Asp	Val	Phe 765	Pro	Pro	Cys 770	Lys
Ile	Ile 770	Asp	Leu	Glu	Ala 775	Val	Lys	Val	Glu	Glu	Glu 780	Leu	Thr	Leu 785	Ser
Trp 785	Thr	Ala	Pro	Gly	Glu 790	Asp	Phe	Asp	Gln	Gly 795	Gln	Ala	Thr	Ser 800	Tyr
Glu	Ile	Arg	Met 805	Ser	Lys	Ser	Leu	Gln	Asn 810	Ile	Gln	Asp	Asp	Phe 815	Asn
Asn	Ala	Ile 820	Leu	Val	Asn	Thr	Ser	Lys 825	Arg	Asn	Pro	Gln	Gln	Ala 830	Gly
Ile	Arg	Glu 835	Ile	Phe	Thr	Phe	Ser 840	Pro	Gln	Ile	Ser 845	Thr	Asn	Gly 850	Pro
Glu	His 850	Gln	Pro	Asn	Gly	Glu	Thr 855	His	Glu	Ser	His 860	Arg	Ile	Tyr 865	Val
Ala 865	Ile	Arg	Ala	Met	Asp 870	Arg	Asn	Ser	Leu	Gln	Ser 875	Ala	Val	Ser 880	Asn
Ile	Ala	Gln	Ala 885	Pro	Leu	Phe	Ile	Pro	Pro 890	Asn	Ser	Asp	Pro	Val 895	Pro

```
<210> 162
<211> 498
<212> DNA
<213> Homo sapien
```

```
<210> 163
<211> 1128
<212> DNA
<213> Homo sapien
```

```
<210> 164
<211> 1310
<212> DNA
<213> Homo sapien
```

<400> 164

```

gggcctggtt cgcaaagaag ctgacttcag agggggaaac tttcttcttt taggaggcgg      60
ttagccctgt tccacgaacc caggagaact gctggccaga ttaattagac attgctatgg      120
gagacgtgta aacacactac ttatcattga tgcataatata aaaccatttt attttcgcta      180
ttatttcaga ggaagcgctt ctgatttggt tcttttttcc ctttttgctc tttctggctg      240
tgtggtttgg agaaagcaca gttggagtag ccggttgcta aataagtcgc gagcgcgagc      300
ggagacgatg cagcggagac tggttcagca gtggagcgtc gcggtgttcc tgctgagcta      360
cgcggtgccc tcctgcgggc gctcgggtga ggggtctcagc cgcgcctca aaagagctgt      420
gtctgaacat cagctcctcc atgacaaggg gaagtccatc caagatttac ggcgacgatt      480
cttccttcac catctgatcg cagaaatcca cacagctgaa atcagagcta cctcggagggt      540
gtcccctaac tccaagccct ctcccaacac aaagaaccac cccgtccgat ttgggtctga      600
tgatgagggc agatacctaa ctcaggaaac taacaagggt gagacgtaca aagagcagcc      660
gctcaagaca cctgggaaga aaaagaaaagg caagcccggg aaacgcaagg agcaggaaaa      720
gaaaaaacgg cgaactcgct ctgcctgggt agactctgga gtgactggga gtgggctaga      780
aggggaccac ctgtctgaca cctccacaac gtcgctggag ctcgattcac ggaggcattg      840
aaattttcag cagagacctt ccaaggacat attgcaggat tctgtaatag tgaacatatg      900
aaaagtatta gaaatattta ttgtctgtaa atactgtaaa tgcattggaa taaaactgtc      960
tccccattg ctctatgaaa ctgcacattg gtcattgtga atattttttt ttttgccaag     1020
gctaattccaa ttattattat cacatttacc ataatttatt ttgtccattg atgtatttat     1080
tttgtaaatg tatcttggtg ctgctgaatt tctatatatt ttgtaacata atgcacttta     1140
gatatacata tcaagtatgt tgataaatga cacaatgaag tgtctctatt ttgtggttga     1200
ttttaatgaa tgccataaata taattatcca aattgatttt cctttgtgcc cgtaaaaaata     1260
acagtatttt aaatttgtaa agaattgtcta ataaaatata atctaattac     1310

```

<210> 165

<211> 177

<212> PRT

<213> Homo sapien

<400> 165

```

Met Gln Arg Arg Leu Val Gln Gln Trp Ser Val Ala Val Phe Leu Leu
1      5      10      15
Ser Tyr Ala Val Pro Ser Cys Gly Arg Ser Val Glu Gly Leu Ser Arg
20     25     30
Arg Leu Lys Arg Ala Val Ser Glu His Gln Leu Leu His Asp Lys Gly
35     40     45
Lys Ser Ile Gln Asp Leu Arg Arg Arg Phe Phe Leu His His Leu Ile
50     55     60
Ala Glu Ile His Thr Ala Glu Ile Arg Ala Thr Ser Glu Val Ser Pro
65     70     75     80
Asn Ser Lys Pro Ser Pro Asn Thr Lys Asn His Pro Val Arg Phe Gly
85     90     95
Ser Asp Asp Glu Gly Arg Tyr Leu Thr Gln Glu Thr Asn Lys Val Glu
100    105    110
Thr Tyr Lys Glu Gln Pro Leu Lys Thr Pro Gly Lys Lys Lys Gly
115    120    125
Lys Pro Gly Lys Arg Lys Glu Gln Glu Lys Lys Lys Arg Arg Thr Arg
130    135    140
Ser Ala Trp Leu Asp Ser Gly Val Thr Gly Ser Gly Leu Glu Gly Asp
145    150    155    160
His Leu Ser Asp Thr Ser Thr Thr Ser Leu Glu Leu Asp Ser Arg Arg
165    170    175
His

```

<210> 166
 <211> 177
 <212> PRT
 <213> Homo sapien

<400> 166

```

Met Gln Arg Arg Leu Val Gln Gln Trp Ser Val Ala Val Phe Leu Leu
 1          5          10          15
Ser Tyr Ala Val Pro Ser Cys Gly Arg Ser Val Glu Gly Leu Ser Arg
 20          25          30
Arg Leu Lys Arg Ala Val Ser Glu His Gln Leu Leu His Asp Lys Gly
 35          40          45
Lys Ser Ile Gln Asp Leu Arg Arg Arg Phe Phe Leu His His Leu Ile
 50          55          60
Ala Glu Ile His Thr Ala Glu Ile Arg Ala Thr Ser Glu Val Ser Pro
 65          70          75          80
Asn Ser Lys Pro Ser Pro Asn Thr Lys Asn His Pro Val Arg Phe Gly
 85          90          95
Ser Asp Asp Glu Gly Arg Tyr Leu Thr Gln Glu Thr Asn Lys Val Glu
 100         105         110
Thr Tyr Lys Glu Gln Pro Leu Lys Thr Pro Gly Lys Lys Lys Lys Gly
 115         120         125
Lys Pro Gly Lys Arg Lys Glu Gln Glu Lys Lys Lys Arg Arg Thr Arg
 130         135         140
Ser Ala Trp Leu Asp Ser Gly Val Thr Gly Ser Gly Leu Glu Gly Asp
 145         150         155         160
His Leu Ser Asp Thr Ser Thr Thr Ser Leu Glu Leu Asp Ser Arg Arg
 165         170         175
His

```

<210> 167
 <211> 3362
 <212> DNA
 <213> Homo sapien

<400> 167

```

cacaatgtat gcagcaggct cagtgtgagt gaactggagg cttctctaca acatgaccca      60
aaggagcatt gcaggtccta ttgcaacct gaagtttgtg actctcctgg ttgccttaag      120
ttcagaactc ccattcctgg gagctggagt acagcttcaa gacaatgggt ataatggatt      180
gtcattgca attaatcctc aggtacctga gaatcagaac ctcatctcaa acattaagga      240
aatgataact gaagcttcat ttacctatt taatgctacc aagagaagag tatttttcag      300
aaatataaag attttaatac ctgccacatg gaaagctaat aataacagca aaataaaaca      360
agaatcatat gaaaaggcaa atgtcatagt gactgactgg tatggggcac atggagatga      420
tcacatacacc ctacaatata gaggggtgtg aaaagagggg aaatacattc atttcacacc      480
taatttccta ctgaatgata acttaacagc tggctacgga tcacgaggcc gagtgtttgt      540
ccatgaatgg gcccacctcc gttggggtgt gttcgatgag tataacaatg acaaaccttt      600
ctacataaat gggcaaaatc aaattaaagt gacaaggtgt tcacttgaca tcacaggcat      660
ttttgtgtgt gaaaaaggtc cttgccccca agaaaactgt attattagta agctttttta      720
agaaggatgc acctttatct acaatagcac ccaaaatgca actgcatcaa taatgttcat      780
gcaaagttta tcttctgtgg ttgaattttg taatgcaagt acccacaacc aagaagcacc      840
aaacctacag aaccagatgt gcagcctcag aagtgcattg gatgtaatca cagactctgc      900
tgactttcac cacagctttc ccatgaacgg gactgagctt ccacctctc ccacattctc      960

```

```

gcttgtagag gctgggtgaca aagtgggtctg tttagtgtctg gatgtgtcca gcaagatggc 1020
agaggctgac agactccttc aactacaaca agcgcgagaa ttttatttga tgcagattgt 1080
tgaaattcat accttcgtgg gcattgccag tttcgacagc aaaggagaga tcagagccca 1140
gctacaccaa attaacagca atgatgatcg aaagtgtctg gtttcataatc tgcccaccac 1200
tgtatcagct aaaacagaca tcagcatttg ttcagggtt aagaaaggat ttgagggtgt 1260
tgaaaaactg aatggaaaag cttatggctc tgtgatgata ttagtgacca gcggagatga 1320
taagcttctt ggcaattgct taccactgt gctcagcagt ggttcaacaa ttcactccat 1380
tgccctgggt tcatctgcag ccccaaactc ggaggaatta tcacgtctta caggaggttt 1440
aaagtctctt gttccagata tatcaaactc caatagcatg attgatgctt tcagtagaat 1500
ttcctctgga actggagaca ttttcagca acatattcag cttgaaagta cagggtgaaa 1560
tgtcaaacct caccatcaat tgaaaaacac agtgactgtg gataaactg tgggcaacga 1620
cactatgttt ctagttacgt ggcaggccag tggctcctct gagattatat tatttgatcc 1680
tgatggacga aaatactaca caaataattt tatcaccaat ctaacttttc ggacagctag 1740
tctttggatt ccaggaacag ctaagcctgg gcaactggact tacaccctga tgtgtttcca 1800
ccatgcaaaa ttattgacct ggaagctgta aaagtagaag aggaattgac cctatcttgg 1860
acagcacctg gagaagactt tgatcagggc caggctacaa gctatgaaat aagaatgagt 1920
aaaagtctac agaatatcca agatgacttt aacaatgcta ttttagtaaa tacatcaaag 1980
cgaaatcctc agcaagctgg catcaggagg atatttacgt tctcacccca aatttccacg 2040
aatggacctg aacatcagcc aaatggagaa acacatgaaat gccacagaat ttatgttgca 2100
atcagagcaa tggataggaa ctctttacag tctgctgtat ctaacattgc ccaggcgcct 2160
ctgtttatct ccccaattc tgatcctgta cctgccagag attatcttat attgaaagga 2220
gttttaacag caatgggttt gataggaatc atttgcctta ttatagtgtt gacacatcat 2280
actttaagca ggaaaaagag agcagacaag aaagagaatg gaacaaaatt attataaata 2340
aatatccaaa gtgtcttctc tcttagatat aagaccatg gccttcgact acaaaaacat 2400
actaacaag tcaaattaac atcaaaactg tattaaaatg cattgagttt ttgtacaata 2460
cagataagat ttttacatgg tagatcaaca aattcttttt gggggtagat tagaaaaccc 2520
ttacactttg gctatgaaca aataataaaa attattcttt aaagtaatgt ctttaaaggc 2580
aaagggaagg gtaagtctg accagtgtca aggaaagttt gttttattga ggtggaaaaa 2640
tagccccaag cagagaaaag gagggtaggt ctgcattata actgtctgtg tgaagcaatc 2700
atttagttac ttgattaat ttttcttttc tcttatctg tgcagaacag gttgcttgtt 2760
tacaactgaa gatcatgcta tatttcataat atgaagcccc taatgcaaag ctctttacct 2820
cttgctattt tgttatatat attacagatg aaatctcact gctaattgctc agagatcttt 2880
tttactgta agaggtaac ttttaacaata tgggtattac ctttgtctct tcataccggt 2940
tttatgacaa aggtctattg aatttatttg tttgtaagtt tctactccca tcaaagcagc 3000
tttctaagtt attgccttgg ttattatgga tgatagttat agcccttata atgccttaac 3060
taaggaagaa aagatgttat tctgagtttg ttttaataca tatatgaaca tatagtttta 3120
ttcaattaaa ccaaagaaga ggtcagcagg gagatactaa cctttggaat tgattagctg 3180
gctctgtttt ttggttaaat aagagtcttt aatcctttct ccatcaagag ttacttacca 3240
agggcagggg aaggggggata tagaggtcac aaggaaataa aaatcatctt tcactcttaa 3300
ttttactcct tctctttatt tttttaaaag attatcgaac aataaaatca tttgcctttt 3360
tt

```

```

<210> 168
<211> 2784
<212> DNA
<213> Homo sapien

```

```

<400> 168
tctgcacca tattgaaaac ctgacacaat gtatgcagca ggctcagtgt gagtgaactg 60
gaggcttctc tacaacatga cccaaaggag cattgcaggt cctatttgca acctgaagtt 120
tgtactctc ctggttgctc taagttcaga actccattc ctgggagctg gactacagct 180
tcaagacaat ggggtataatg gattgctcat tgcaattaa cctcaggtac ctgagaatca 240
gaacctcatc tcaaacatta aggaaatgat aactgaagct tcattttacc tatttaatgc 300
taccaagaga agagtatttt tcagaaatat aaagatttta atacctgcca catggaaagc 360

```

```

taataataac agcaaaataa aacaagaatc atatgaaaag gcaaattgtca tagtgactga 420
ctgggtatggg gcacatggag atgatccata caccctacaa tacagagggg gtggaaaaga 480
gggaaaatac attcattttca cacctaattt cctactgaat gataacttaa cagctggcta 540
cggatcacga ggccgagtggt ttgtccatga atgggcccac ctccgttggg gtgtgttcga 600
tgagtataac aatgacaaac ctttctacat aaatgggcaa aatcaaatta aagtgacaag 660
gtgttcacat gacatcacag gcatttttgt gtgtgaaaaa ggtccttgcc cccaagaaaa 720
ctgtattatt agtaagcttt ttaaagaagg atgcaccttt atctacaata gcacccaaaa 780
tgcaactgca tcaataatgt tcatgcaaag tttatcttct gtggttgaat tttgtaatgc 840
aagtaccacac aaccaagaag caccaaacct acagaaccag atgtgcagcc tcagaagtgc 900
atgggatgta atcacagact ctgctgactt tcaccacagc tttcccatga acgggactga 960
gcttccacct cctcccacat tctcgttgt agaggctggg gacaaagtgg tctgtttagt 1020
gctggatgtg tccagcaaga tggcagaggc tgacagactc cttcaactac aacaagccgc 1080
agaattttat ttgatgcaga ttgttgaaat tcataccttc gtgggcattg ccagttttcga 1140
cagcaaagga gagatcagag ccagctaca ccaaattaac agcaatgatg atcgaaagt 1200
gctggtttca tatctgcca cactgtatc agctaaaaca gacatcagca tttgttcagg 1260
gcttaagaaa ggatttgagg tgggtgaaaa actgaatgga aaagcttatg gctctgtgat 1320
gatattagtg accagcggag atgataagct tcttggcaat tgcttaccga ctgtgctcag 1380
cagtgggtca acaattcact ccattgcctt ggggttcact gcagcccaa atctggagga 1440
attatcacgt cttacaggag gtttaaagtt ctttgttcca gatatatcaa actccaatag 1500
catgattgat gctttcagta gaatttcctc tggaaactgga gacattttcc agcaacatat 1560
tcagcttgaa agtacagggtg aaaatgtcaa acctcaccat caattgaaaa acacagtgc 1620
tgtggataat actgtgggca acgacactat gtttctagtt acgtggcagg ccagtggtcc 1680
tctgagatt atattatttg atcctgatgg acgaaaatac tacacaaata attttatcac 1740
caatctaact tttcggacag ctagtctttg gattccagga acagctaagc ctgggcaactg 1800
gacttacacc ctgaacaata cccatcattc tctgcaagcc ctgaaagtga cagtgcctc 1860
tcgcgcctcc aactcagctg tgcctccagc cactgtggaa gcctttgtgg aaagagacag 1920
cctccatttt cctcatcctg tgatgattta tgccaatgtg aaacagggat tttatcccat 1980
tcttaatgcc actgtcactg ccacagttga gccagagact ggagatcctg ttacgctgag 2040
actccttgat gatggagcag gtgctgatgt tataaaaaat gatggaattt actcgaggta 2100
ttttttctcc tttgctgcaa atggtagata tagcttgaaa gtgcatgtca atcactctcc 2160
cagcataagc accccagccc actctatttc agggagtcac gctatgtatg taccagggtta 2220
cacagcaaac ggtaatatcc agatgaatgc tccaaggaaa tcagtaggca gaaatgagga 2280
ggagcgaaaag tggggcttta gccgagtcag ctcaggaggc tccttttcag tgctgggagt 2340
tccagctggc cccaccctg atgtgtttcc accatgcaaa attattgacc tggaagctgt 2400
aaatagaaga ggaattgacc ctatcttgga cagcacctgg agaagacttt gatcagggcc 2460
aggctacaag ctatgaaata agaattgagta aaagtctaca gaatatccaa gatgacttta 2520
acaatgctat tttagtaaat acatcaaagc gaaatcctca gcaagctggc atcagggaga 2580
tattttacgtt ctcaccccaa atttccacga atggacctga acatcagcca aatggagaaa 2640
cacatgaaag ccacagaatt tatgttgcaa tacgagcaat ggataggaac tccttacagt 2700
ctgctgtatc taacattgcc caggcgctc tgtttattcc cccaattct gatcctgtac 2760
ctgccagaga ttatcttata ttga 2784

```

<210> 169

<211> 592

<212> PRT

<213> Homo sapien

<400> 169

```

Met Thr Gln Arg Ser Ile Ala Gly Pro Ile Cys Asn Leu Lys Phe Val
1           5           10           15
Thr Leu Leu Val Ala Leu Ser Ser Glu Leu Pro Phe Leu Gly Ala Gly
20           25           30
Val Gln Leu Gln Asp Asn Gly Tyr Asn Gly Leu Leu Ile Ala Ile Asn
35           40           45

```


Pro	Gln	Val	Pro	Glu	Asn	Gln	Asn	Leu	Ile	Ser	Asn	Ile	Lys	Glu	Met
50						55					60				
Ile	Thr	Glu	Ala	Ser	Phe	Tyr	Leu	Phe	Asn	Ala	Thr	Lys	Arg	Arg	Val
65					70					75					80
Phe	Phe	Arg	Asn	Ile	Lys	Ile	Leu	Ile	Pro	Ala	Thr	Trp	Lys	Ala	Asn
				85					90					95	
Asn	Asn	Ser	Lys	Ile	Lys	Gln	Glu	Ser	Tyr	Glu	Lys	Ala	Asn	Val	Ile
			100					105					110		
Val	Thr	Asp	Trp	Tyr	Gly	Ala	His	Gly	Asp	Asp	Pro	Tyr	Thr	Leu	Gln
		115				120						125			
Tyr	Arg	Gly	Cys	Gly	Lys	Glu	Gly	Lys	Tyr	Ile	His	Phe	Thr	Pro	Asn
	130					135				140					
Phe	Leu	Leu	Asn	Asp	Asn	Leu	Thr	Ala	Gly	Tyr	Gly	Ser	Arg	Gly	Arg
145				150						155					160
Val	Phe	Val	His	Glu	Trp	Ala	His	Leu	Arg	Trp	Gly	Val	Phe	Asp	Glu
				165					170					175	
Tyr	Asn	Asn	Asp	Lys	Pro	Phe	Tyr	Ile	Asn	Gly	Gln	Asn	Gln	Ile	Lys
			180					185					190		
Val	Thr	Arg	Cys	Ser	Ser	Asp	Ile	Thr	Gly	Ile	Phe	Val	Cys	Glu	Lys
		195				200						205			
Gly	Pro	Cys	Pro	Gln	Glu	Asn	Cys	Ile	Ile	Ser	Lys	Leu	Phe	Lys	Glu
	210					215					220				
Gly	Cys	Thr	Phe	Ile	Tyr	Asn	Ser	Thr	Gln	Asn	Ala	Thr	Ala	Ser	Ile
225					230					235					240
Met	Phe	Met	Gln	Ser	Leu	Ser	Ser	Val	Val	Glu	Phe	Cys	Asn	Ala	Ser
				245						250				255	
Thr	His	Asn	Gln	Glu	Ala	Pro	Asn	Leu	Gln	Asn	Gln	Met	Cys	Ser	Leu
			260					265					270		
Arg	Ser	Ala	Trp	Asp	Val	Ile	Thr	Asp	Ser	Ala	Asp	Phe	His	His	Ser
		275					280					285			
Phe	Pro	Met	Asn	Gly	Thr	Glu	Leu	Pro	Pro	Pro	Pro	Thr	Phe	Ser	Leu
	290					295					300				
Val	Glu	Ala	Gly	Asp	Lys	Val	Val	Cys	Leu	Val	Leu	Asp	Val	Ser	Ser
305					310					315					320
Lys	Met	Ala	Glu	Ala	Asp	Arg	Leu	Leu	Gln	Leu	Gln	Gln	Ala	Ala	Glu
				325					330					335	
Phe	Tyr	Leu	Met	Gln	Ile	Val	Glu	Ile	His	Thr	Phe	Val	Gly	Ile	Ala
			340					345					350		
Ser	Phe	Asp	Ser	Lys	Gly	Glu	Ile	Arg	Ala	Gln	Leu	His	Gln	Ile	Asn
		355					360					365			
Ser	Asn	Asp	Asp	Arg	Lys	Leu	Leu	Val	Ser	Tyr	Leu	Pro	Thr	Thr	Val
	370					375					380				
Ser	Ala	Lys	Thr	Asp	Ile	Ser	Ile	Cys	Ser	Gly	Leu	Lys	Lys	Gly	Phe
385					390					395					400
Glu	Val	Val	Glu	Lys											

			260					265				270			
Arg	Ser	Ala	Trp	Asp	Val	Ile	Thr	Asp	Ser	Ala	Asp	Phe	His	His	Ser
		275					280					285			
Phe	Pro	Met	Asn	Gly	Thr	Glu	Leu	Pro	Pro	Pro	Pro	Thr	Phe	Ser	Leu
	290					295					300				
Val	Glu	Ala	Gly	Asp	Lys	Val	Val	Cys	Leu	Val	Leu	Asp	Val	Ser	Ser
305					310					315					320
Lys	Met	Ala	Glu	Ala	Asp	Arg	Leu	Leu	Gln	Leu	Gln	Gln	Ala	Ala	Glu
				325					330					335	
Phe	Tyr	Leu	Met	Gln	Ile	Val	Glu	Ile	His	Thr	Phe	Val	Gly	Ile	Ala
			340					345					350		
Ser	Phe	Asp	Ser	Lys	Gly	Glu	Ile	Arg	Ala	Gln	Leu	His	Gln	Ile	Asn
		355					360					365			
Ser	Asn	Asp	Asp	Arg	Lys	Leu	Leu	Val	Ser	Tyr	Leu	Pro	Thr	Thr	Val
	370					375					380				
Ser	Ala	Lys	Thr	Asp	Ile	Ser	Ile	Cys	Ser	Gly	Leu	Lys	Lys	Gly	Phe
385					390					395					400
Glu	Val	Val	Glu	Lys	Leu	Asn	Gly	Lys	Ala	Tyr	Gly	Ser	Val	Met	Ile
				405					410					415	
Leu	Val	Thr	Ser	Gly	Asp	Asp	Lys	Leu	Leu	Gly	Asn	Cys	Leu	Pro	Thr
			420					425					430		
Val	Leu	Ser	Ser	Gly	Ser	Thr	Ile	His	Ser	Ile	Ala	Leu	Gly	Ser	Ser
		435					440					445			
Ala	Ala	Pro	Asn	Leu	Glu	Glu	Leu	Ser	Arg	Leu	Thr	Gly	Gly	Leu	Lys
	450					455					460				
Phe	Phe	Val	Pro	Asp	Ile	Ser	Asn	Ser	Asn	Ser	Met	Ile	Asp	Ala	Phe
465					470					475					480
Ser	Arg	Ile	Ser	Ser	Gly	Thr	Gly	Asp	Ile	Phe	Gln	Gln	His	Ile	Gln
				485				490						495	
Leu	Glu	Ser	Thr	Gly	Glu	Asn	Val	Lys	Pro	His	His	Gln	Leu	Lys	Asn
			500					505					510		
Thr	Val	Thr	Val	Asp	Asn	Thr	Val	Gly	Asn	Asp	Thr	Met	Phe	Leu	Val
		515					520					525			
Thr	Trp	Gln	Ala	Ser	Gly	Pro	Pro	Glu	Ile	Ile	Leu	Phe	Asp	Pro	Asp
	530					535					540				
Gly	Arg	Lys	Tyr	Tyr	Thr	Asn	Asn	Phe	Ile	Thr	Asn	Leu	Thr	Phe	Arg
545					550					555					560
Thr	Ala	Ser	Leu	Trp	Ile	Pro	Gly	Thr	Ala	Lys	Pro	Gly	His	Trp	Thr
				565					570					575	
Tyr	Thr	Leu	Asn	Asn	Thr	His	His	Ser	Leu	Gln	Ala	Leu	Lys	Val	Thr
			580					585					590		
Val	Thr	Ser	Arg	Ala	Ser	Asn	Ser	Ala	Val	Pro	Pro	Ala	Thr	Val	Glu
		595					600					605			
Ala	Phe	Val	Glu	Arg	Asp	Ser	Leu	His	Phe	Pro	His	Pro	Val	Met	Ile
	610														

690	695	700
Pro Gly Ser His Ala Met Tyr Val Pro Gly Tyr Thr Ala Asn Gly Asn		
705	710	715
Ile Gln Met Asn Ala Pro Arg Lys Ser Val Gly Arg Asn Glu Glu Glu		720
	725	730
Arg Lys Trp Gly Phe Ser Arg Val Ser Ser Gly Gly Ser Phe Ser Val		735
	740	745
Leu Gly Val Pro Ala Gly Pro His Pro Asp Val Phe Pro Pro Cys Lys		750
	755	760
Ile Ile Asp Leu Glu Ala Val Asn Arg Arg Gly Ile Asp Pro Ile Leu		765
	770	775
Asp Ser Thr Trp Arg Arg Leu		780
785	790	

<210> 171

<211> 1491

<212> DNA

<213> Homo sapien

<400> 171

cctcctgcc	gccaaagtga	gacatgctta	cttcccocttc	accttccttc	atgatgtggg	60
aagagtgtcg	caaccagcc	ctagccaacg	cgcagtgaga	gggagtgtgc	cgagggttc	120
tgagaagggt	tctctcacat	ctagaaagaa	gcgttaaga	tgtggcagcc	cctcttcttc	180
aagtggctct	tgctctgttg	ccctgggagt	tctcaaattg	ctgcagcagc	ctccaccag	240
cctgaggatg	acatcaatac	acagaggaag	aagagtcagg	aaaagatgag	agaagttaca	300
gactctcctg	ggcgaccccg	agagcttacc	attcctcaga	cttcttcaca	tggtgctaac	360
agatttggtc	ctaaaagtaa	agctctagag	gocgtcaa	atggaataga	agccgggttc	420
caccatattg	attctgcaca	tggtttaca	aatgaggagc	aggttggtg	ggccatccga	480
agcaagattg	cagatggcag	tgtgaagaga	gaagacatat	tctacacttc	aaagctttgg	540
agcaattccc	atcgaccaga	gttggtccga	ccagccttgg	aaaggtcact	gaaaaatctt	600
caattggact	atgttgacct	ctatcttatt	cattttccag	tgtctgtaaa	gccaggtgag	660
gaagtgatcc	caaaagatga	aaatggaaaa	atactatttg	acacagtggg	tctctgtgcc	720
acatgggagg	ccatggagaa	gtgtaaagat	gcaggattgg	ccaagtccat	cgggggtgtc	780
aacttcaacc	acaggctgct	ggagatgatc	ctcaacaagc	cagggtctca	gtacaagcct	840
gtctgcaacc	aggtggaatg	tcatecttac	ttcaaccaga	gaaaactgct	ggattttctg	900
aagtcaaaag	acattgttct	ggttgcctat	agtgtctctg	gatcccatcg	agaagaacca	960
tggttggtg	cgaactcccc	ggtgtctctg	gaggaccag	tcctttgtgc	cttggcaaaa	1020
aagcacaagc	gaaccccagc	cctgattgcc	ctgcgtctac	agctgcagcg	tggggttggtg	1080
gtcctggcca	agagctacaa	tgagcagcgc	atcagacaga	acgtgcaggt	gtttgaattc	1140
cagttgactt	cagaggagat	gaaagccata	gatggcctaa	acagaaatgt	gcgatatttg	1200
acccttgata	tttttgctgg	ccccctaat	tatccatttt	ctgatgaata	ttaacatgga	1260
gggcattgca	tgaggtctgc	cagaaggccc	tgcgtgtgga	tggtgacaca	gaggatggct	1320
ctatgctggt	gactggacac	atgcctctctg	gttaaattctc	tcctgcttgg	cgacttcagt	1380
aagctacagc	taagcccatc	ggccggaaaa	gaaagacaat	aattttgttt	ttcattttga	1440
aaaaattaaa	tgctctctcc	taaagattct	tcacctaaaa	aaaaaaaaa	a	1491

<210> 172

<211> 364

<212> PRT

<213> Homo sapien

<400> 172

Met Trp Gln Pro Leu Phe Phe Lys Trp Leu Leu Ser Cys Cys Pro Gly
1 5 10 15

Ser Ser Gln Ile Ala Ala Ala Ala Ser Thr Gln Pro Glu Asp Asp Ile
 20 25 30
 Asn Thr Gln Arg Lys Lys Ser Gln Glu Lys Met Arg Glu Val Thr Asp
 35 40 45
 Ser Pro Gly Arg Pro Arg Glu Leu Thr Ile Pro Gln Thr Ser Ser His
 50 55 60
 Gly Ala Asn Arg Phe Val Pro Lys Ser Lys Ala Leu Glu Ala Val Lys
 65 70 75 80
 Leu Ala Ile Glu Ala Gly Phe His His Ile Asp Ser Ala His Val Tyr
 85 90 95
 Asn Asn Glu Glu Gln Val Gly Leu Ala Ile Arg Ser Lys Ile Ala Asp
 100 105 110
 Gly Ser Val Lys Arg Glu Asp Ile Phe Tyr Thr Ser Lys Leu Trp Ser
 115 120 125
 Asn Ser His Arg Pro Glu Leu Val Arg Pro Ala Leu Glu Arg Ser Leu
 130 135 140
 Lys Asn Leu Gln Leu Asp Tyr Val Asp Leu Tyr Leu Ile His Phe Pro
 145 150 155 160
 Val Ser Val Lys Pro Gly Glu Glu Val Ile Pro Lys Asp Glu Asn Gly
 165 170 175
 Lys Ile Leu Phe Asp Thr Val Asp Leu Cys Ala Thr Trp Glu Ala Met
 180 185 190
 Glu Lys Cys Lys Asp Ala Gly Leu Ala Lys Ser Ile Gly Val Ser Asn
 195 200 205
 Phe Asn His Arg Leu Leu Glu Met Ile Leu Asn Lys Pro Gly Leu Lys
 210 215 220
 Tyr Lys Pro Val Cys Asn Gln Val Glu Cys His Pro Tyr Phe Asn Gln
 225 230 235 240
 Arg Lys Leu Leu Asp Phe Cys Lys Ser Lys Asp Ile Val Leu Val Ala
 245 250 255
 Tyr Ser Ala Leu Gly Ser His Arg Glu Glu Pro Trp Val Asp Pro Asn
 260 265 270
 Ser Pro Val Leu Leu Glu Asp Pro Val Leu Cys Ala Leu Ala Lys Lys
 275 280 285
 His Lys Arg Thr Pro Ala Leu Ile Ala Leu Arg Tyr Gln Leu Gln Arg
 290 295 300
 Gly Val Val Val Leu Ala Lys Ser Tyr Asn Glu Gln Arg Ile Arg Gln
 305 310 315 320
 Asn Val Gln Val Phe Glu Phe Gln Leu Thr Ser Glu Glu Met Lys Ala
 325 330 335
 Ile Asp Gly Leu Asn Arg Asn Val Arg Tyr Leu Thr Leu Asp Ile Phe
 340 345 350
 Ala Gly Pro Pro Asn Tyr Pro Phe Ser Asp Glu Tyr
 355 360

<210> 173

<211> 1988

<212> DNA

<213> Homo sapiens

<400> 173

cgggagcgcg ctccccgcgg cctcttcgct tttgtggcgg cgccccgcgt cgcaggccac 60
 tctctgtgtg cgccgctccc gcgcgtcct ccgacccgct ccgctccgct ccgctcggcc 120

```

ccgcgcgcgc cgtcaacatg atccgctgcg gacctggcctg cgagcgcctgc cgctggatcc 180
tgcacctgct cctactcagc gccatcgcc tgcacatcat cgcgcctggcc ggccgcggct 240
gggtgcagtc tagcgaccac ggccagacgt cctcgctgtg gtggaaatgc tcccaagagg 300
gcggcggcag cgggtcctac gaggagggtc gtcagagcct catggagtag gcgtggggta 360
gagcagcggc tgccatgctc ttctgtggct tcatcatcct ggtgatctgt ttcacctctc 420
ccttcttcgc cctctgtgga cccagatgc ttgtcttctc gagagtgatt ggaggtctcc 480
ttgccttggc tgctgtgttc cagatcatct ccctggtaat ttaccccgct aagtacaccc 540
agaccttcac ccttcatgcc aacctgctg tcaactacat ctataactgg gcctacggct 600
ttgggtgggc agccacgatt atcctgatcg gctgtgcctt cttcttctgc tgccctcccca 660
actacgaaga tgaccttctg ggcaatgcc agcccaggta cttctacaca tctgcctaac 720
ttgggaatga atgtgggaga aaatcgctgc tgctgagatg gactccagaa gaagaaactg 780
tttctccagg cgactttgaa cccatttttt ggcagtgttc atattattaa actagtcaaa 840
aatgctaaaa taatttgga gaaaatattt tttaagtagt gttatagttt catgtttatc 900
ttttattatg ttttgtgaag ttgtgtcttt tcaactaatta cctatactat gccaatattt 960
ccttatatct atccataaca ttatactac atttgtaaga gaatatgcac gtgaaactta 1020
acactttata aggtaaaaat gaggtttcca agatttaata atctgatcaa gttcttgta 1080
tttccaaata gaatggactt ggtctgttaa gggctaagga gaagaggaag ataagggtta 1140
aagttgttaa tgaccaaaaca ttctaaaaga aatgcaaaaa aaaagtttat tttcaagcct 1200
tcgaactatt taaggaaagc aaaatcattt cctaaatgca tatcatttgt gagaatttct 1260
cattaatatt ctgaatcatt catttcagct aaggcttcat gttgactcga tatgtcatct 1320
aggaaagtag tatttcattg tccaaacctg ttgccatagt tggtaaggct ttcctttaag 1380
tgtgaaatat ttagatgaaa ttttctcttt taaagttctt tatagggtta ggggtgtgga 1440
aaatgctata ttaataaatc tgtagtgttt tgtgtttata tgttcagaac cagagtagac 1500
tggaattgaaa gatggactgg gtctaattta tcatgactga tagatctggt taagttgtgt 1560
agtaaagcat taggagggtc attcytgtca caaagtgcc actaaaacag cctcaggaga 1620
ataaatgact tgcttttcta aatctcaggt ttatctgggc tctatcatat agacaggctt 1680
ctgatagttt gcarctgtaa gcagaaacct acatatagtt aaaatcctgg tctttcttgg 1740
taaacagatt ttaaattgtc gatataaaac atgccacagg agaattcggg gatttgagtt 1800
tctctgaata gcatatatat gatgcacagg ataggtcatt atgatttttt accatttcca 1860
cttacataat gaaaaccaat tcatttttaa tatcagatta ttattttgta agttgtggaa 1920
aaagctaatt gtagttttca ttatgaagtt ttcccaataa accagggtatt ctaaaaaaaa 1980
aaaaaaaa

```

<210> 174

<211> 238

<212> PRT

<213> Homo sapiens

<400> 174

Gly Ala Ala Ser Pro Arg Pro Leu Arg Phe Cys Gly Gly Ala Arg Ala
5 10 15

Arg Arg Pro Leu Ser Ala Val Ala Arg Pro Ala Arg Ser Ser Asp Pro
20 25 30

Leu Arg Ser Ala Pro Leu Gly Pro Ala Pro Pro Val Asn Met Ile Arg
35 40 45

Cys Gly Leu Ala Cys Glu Arg Cys Arg Trp Ile Leu Pro Leu Leu Leu
50 55 60

Leu Ser Ala Ile Ala Phe Asp Ile Ile Ala Leu Ala Gly Arg Gly Trp
65 70 75 80

```

<210> 175
<211> 4181
<212> DNA
<213> Homo sapiens

<220>
<221> unsure
<222> (3347)
<223> n=A,T,C or G
<221> unsure
<222> (3502)
<223> n=A,T,C or G
<221> unsure
<222> (3506)
<223> n=A,T,C or G
<221> unsure
<222> (3520)
<223> n=A,T,C or G
<221> unsure
<222> (3538)
<223> n=A,T,C or G
<221> unsure

```

<222> (3549)
 <223> n=A,T,C or G
 <221> unsure
 <222> (3646)
 <223> n=A,T,C or G
 <221> unsure
 <222> (3940)
 <223> n=A,T,C or G
 <221> unsure
 <222> (3968)
 <223> n=A,T,C or G
 <221> unsure
 <222> (3974)
 <223> n=A,T,C or G
 <221> unsure
 <222> (4036)
 <223> n=A,T,C or G
 <221> unsure
 <222> (4056)
 <223> n=A,T,C or G
 <221> unsure
 <222> (4062)
 <223> n=A,T,C or G
 <221> unsure
 <222> (4080)
 <223> n=A,T,C or G
 <221> unsure
 <222> (4088)
 <223> n=A,T,C or G
 <221> unsure
 <222> (4115)
 <223> n=A,T,C or G

<400> 175

ggtggtgctg	tttgggtgtg	agctaggctt	tttcttttct	ttctctttta	aaacacatct	60
agacaaggaa	aaaacaagcc	tcggatctga	tttttcactc	ctcgttcttg	tgcttggttc	120
ttactgtgtt	tgtgtatttt	aaaggcgaga	agacgagggg	aacaaaacca	gctggatcca	180
tccatcaccg	tgggtggttt	taatttttctg	ttttttctcg	ttattttttt	ttaaacaacc	240
actcttcaca	atgaacaaac	tgtatatcgg	aaacctcagc	gagaacgccg	ccccctcgga	300
cctagaaaagt	atcttcaagg	acgccaagat	cccggtgtcg	ggacccttcc	tggtgaagac	360
tggctacgcg	ttcgtggact	gcccggacga	gagctgggcc	ctcaaggcca	tcgaggcgct	420
ttcaggtaaa	atagaactgc	acgggaaacc	catagaagtt	gagcactcgg	tcccaaaaag	480
gcaaaggatt	cggaacttcc	agatacgaaa	tatccgcctt	catttacagt	gggaggtgct	540
ggatagttaa	ctagtccagt	atggagtggg	ggagagctgt	gagcaagtga	acactgactc	600
ggaaactgca	gttgtaaata	taacctattc	cagtaaggac	caagctagac	aagcactaga	660
caaactgaat	ggatttcagt	tagagaattt	caccttgaaa	gtagcctata	tccctgatga	720
aatggccgcc	cagcaaaaacc	ccttgcagca	gccccgaggt	cgccgggggc	ttgggcagag	780
gggtctctca	aggcaggggt	ctccaggatc	cgtatccaag	cagaaaccat	gtgatttgcc	840
tctgcgctg	ctgggtccca	cccaatttgt	tggagccatc	ataggaaaag	aaggtgccac	900
cattcggaac	atcaccaaac	agacccagtc	taaaatcgat	gtccaccgta	aagaaaatgc	960
gggggctgct	gagaagtoga	ttactatcct	ctctactcct	gaaggcacct	ctgcggcttg	1020
taagcttatt	ctggagatta	tgcataagga	agctcaagat	ataaaattca	cagaagagat	1080
ccccttgaag	atttttagctc	ataataactt	tgttggacgt	cttattggta	aagaaggaag	1140
aaatcttaaa	aaaattgagc	aagacacaga	cactaaaatc	acgatatctc	cattgcagga	1200


```

attgacgctg tataatccag aacgcactat tacagttaaa ggcaatgttg agacatgtgc 1260
caaagctgag gaggagatca tgaagaaaat cagggagtct tatgaaaatg atattgcttc 1320
tatgaatctt caagcacatt taattcctgg attaaatctg aacgccttgg gtctgttccc 1380
acccaacttca gggatgccac ctcccacctc agggcccccct tcagccatga ctccctcccta 1440
cccgagttt gagcaatcag aaacgggagac tgttcatcag tttatcccag ctctatcagt 1500
cggtgccatc atcggaagc agggccagca catcaagcag ctttctcgct ttgctggagc 1560
ttcaattaag attgtctcag cggaagcacc agatgctaaa gtgaggatgg tgattatcac 1620
tggaaccacca gaggtcagc tcaaggctca gggaagaatt tatggaaaaa ttaaagaaga 1680
aaactttgtt agtcctaaag aagaggtgaa acttgaagct catatcagag tgccatcctt 1740
tgetgctggc agagttattg gaaaaggagg caaacgggtg aatgaacttc agaatttgtc 1800
aagtgcagaa gttgttgctc ctctgaccca gacacctgat gagaatgacc aagtggttgt 1860
caaaataact ggtcacttct atgcttgcca gggtgcccag agaaaaattc aggaaattct 1920
gactcaggtt aagcagcacc aacaacagaa ggctctgcaa agtgaggacc ctcagtcaag 1980
acggaagtaa aggtcagga aacagcccac cacagaggca gatgccaaac caaagacaga 2040
ttgcttaacc aacagatggg cgtgaccccc ctatccagaa tcacatgcac aagtttttac 2100
ctagccagtt gtttctgagg accaggcaac ttttgaactc ctgtctctgt gagaatgtat 2160
actttatgct ctctgaaatg tatgacaccc agctttaaaa caaacaacaa aacaaacaaa 2220
aaaagggtgg gggaggagg gaaagagaag agctctgcac ttccctttgt tgtagtctca 2280
cagtataaca gatattctaa ttcttcttaa tttcccccga taatgccaga aattggctta 2340
atgatgcttt cactaaattc atcaaataga ttgctcctaa atccaattgt taaaattgga 2400
tcagaataat tatcacagga acttaaatgt taagccatta gcatagaaaa actgttctca 2460
gttttatttt tacctaacac taacatgagt aacctagggt aagtgtgaa tgggtgttggc 2520
aggggtatta aacgtgcatt tttactcaac tacctcaggt attcagtaat acaatgaaaa 2580
gcaaaattgt tctttttttt tgaaaatttt atatacttta taatgataga agtccaaccg 2640
tttttttaaa aataaattta aaatttaaca gcaatcagct aacaggcaaa ttaagatttt 2700
tacttctggc tgggtgacagt aaagctggaa aattaatttc aggggttttt gaggcttttg 2760
acacagttat tagttaaatc aaatgttcaa aaatacggag cagtgcctag tatctggaga 2820
gcagcactac catttattct ttcatattata gttgggaaag tttttgacgg tactaacaaa 2880
gtggctgcag gagatttttg aacggctggt ttaaatggct tcaggagact tcagtttttt 2940
gttttagctac atgattgaat gcataataaa tgctttgtgc ttctgactat caatacctaa 3000
agaaagtgca tcagtgaaga gatgcaagac tttcaactga ctggcaaaaa gcaagcttta 3060
gcttgtctta taggatgctt agtttgccac tacacttcag accaatggga cagtcataga 3120
tggtgtgaca gtgtttaaac gcaacaaaag gctacatttc catggggcca gcactgtcat 3180
gagcctcact aagctatttt gaagattttt aagcactgat aaattaaaaa aaaaaaaaaa 3240
aaattagact ccaccttaag tagtaaagta taacaggatt tctgtatact gtgcaatcag 3300
ttctttgaaa aaaaagtcaa aagatagaga atacaagaaa agttttnggg atataatttg 3360
aatgactgtg aaaacatatg acctttgata acgaactcat ttgctcactc cttgacagca 3420
aagcccagta cgtacaattg tgttgggtgt ggggtggtctc caaggccacg ctgctctctg 3480
aattgatttt ttgagttttg gnttgnaaga tgatcacagn catgttacac tgatcttnaa 3540
ggacatatnt tataaccctt taaaaaaaaa atccccctgc tcattcttat ttcgagatga 3600
atttcgatac agactagatg tctttctgaa gatcaattag acattntgaa aatgatttaa 3660
agtgttttcc ttaatgttct ctgaaaacaa gtttcttttg tagttttaac caaaaaagtg 3720
ccctttttgt cactggtttc tcctagcatt catgattttt ttttcacaca atgaattaaa 3780
attgctaaaa tcatggactg gctttctggt tggatttcag gtaagatgtg ttttaaggcca 3840
gagcttttct cagtatttga ttttttccc caatatattga ttttttaaaa atatacacat 3900
aggagctgca tttaaaacct gctggtttta attctgtcan atttcacttc tagcctttta 3960
gtatggcnaa tcanaattta cttttactta agcatttgta atttgagta tctggtacta 4020
gctaagaaat aattcnataa ttgagttttg tactenocaa anatgggtca ttcctcatgn 4080
ataatgtnc cccaatgcag cttcattttc caganacctt gacgcaggat aaattttttc 4140
atcatttagg tccccaaaaa aaaaaaaaaa aaaaaaaaaa a 4181

```

<210> 176

<211> 579

<212> PRT

<213> Homo sapiens

<400> 176

```

Met Asn Lys Leu Tyr Ile Gly Asn Leu Ser Glu Asn Ala Ala Pro Ser
      5                      10                      15

Asp Leu Glu Ser Ile Phe Lys Asp Ala Lys Ile Pro Val Ser Gly Pro
      20                      25                      30

Phe Leu Val Lys Thr Gly Tyr Ala Phe Val Asp Cys Pro Asp Glu Ser
      35                      40                      45

Trp Ala Leu Lys Ala Ile Glu Ala Leu Ser Gly Lys Ile Glu Leu His
      50                      55                      60

Gly Lys Pro Ile Glu Val Glu His Ser Val Pro Lys Arg Gln Arg Ile
      65                      70                      75                      80

Arg Lys Leu Gln Ile Arg Asn Ile Pro Pro His Leu Gln Trp Glu Val
      85                      90                      95

Leu Asp Ser Leu Leu Val Gln Tyr Gly Val Val Glu Ser Cys Glu Gln
      100                     105                     110

Val Asn Thr Asp Ser Glu Thr Ala Val Val Asn Val Thr Tyr Ser Ser
      115                     120                     125

Lys Asp Gln Ala Arg Gln Ala Leu Asp Lys Leu Asn Gly Phe Gln Leu
      130                     135                     140

Glu Asn Phe Thr Leu Lys Val Ala Tyr Ile Pro Asp Glu Met Ala Ala
      145                     150                     155                     160

Gln Gln Asn Pro Leu Gln Gln Pro Arg Gly Arg Arg Gly Leu Gly Gln
      165                     170                     175

Arg Gly Ser Ser Arg Gln Gly Ser Pro Gly Ser Val Ser Lys Gln Lys
      180                     185                     190

Pro Cys Asp Leu Pro Leu Arg Leu Leu Val Pro Thr Gln Phe Val Gly
      195                     200                     205

Ala Ile Ile Gly Lys Glu Gly Ala Thr Ile Arg Asn Ile Thr Lys Gln
      210                     215                     220

Thr Gln Ser Lys Ile Asp Val His Arg Lys Glu Asn Ala Gly Ala Ala
      225                     230                     235                     240

Glu Lys Ser Ile Thr Ile Leu Ser Thr Pro Glu Gly Thr Ser Ala Ala
      245                     250                     255

Cys Lys Ser Ile Leu Glu Ile Met His Lys Glu Ala Gln Asp Ile Lys
      260                     265                     270

```

Phe Thr Glu Glu Ile Pro Leu Lys Ile Leu Ala His Asn Asn Phe Val
 275 280 285
 Gly Arg Leu Ile Gly Lys Glu Gly Arg Asn Leu Lys Lys Ile Glu Gln
 290 295 300
 Asp Thr Asp Thr Lys Ile Thr Ile Ser Pro Leu Gln Glu Leu Thr Leu
 305 310 315 320
 Tyr Asn Pro Glu Arg Thr Ile Thr Val Lys Gly Asn Val Glu Thr Cys
 325 330 335
 Ala Lys Ala Glu Glu Glu Ile Met Lys Lys Ile Arg Glu Ser Tyr Glu
 340 345 350
 Asn Asp Ile Ala Ser Met Asn Leu Gln Ala His Leu Ile Pro Gly Leu
 355 360 365
 Asn Leu Asn Ala Leu Gly Leu Phe Pro Pro Thr Ser Gly Met Pro Pro
 370 375 380
 Pro Thr Ser Gly Pro Pro Ser Ala Met Thr Pro Pro Tyr Pro Gln Phe
 385 390 395 400
 Glu Gln Ser Glu Thr Glu Thr Val His Gln Phe Ile Pro Ala Leu Ser
 405 410 415
 Val Gly Ala Ile Ile Gly Lys Gln Gly Gln His Ile Lys Gln Leu Ser
 420 425 430
 Arg Phe Ala Gly Ala Ser Ile Lys Ile Ala Pro Ala Glu Ala Pro Asp
 435 440 445
 Ala Lys Val Arg Met Val Ile Ile Thr Gly Pro Pro Glu Ala Gln Phe
 450 455 460
 Lys Ala Gln Gly Arg Ile Tyr Gly Lys Ile Lys Glu Glu Asn Phe Val
 465 470 475 480
 Ser Pro Lys Glu Glu Val Lys Leu Glu Ala His Ile Arg Val Pro Ser
 485 490 495
 Phe Ala Ala Gly Arg Val Ile Gly Lys Gly Gly Lys Thr Val Asn Glu
 500 505 510
 Leu Gln Asn Leu Ser Ser Ala Glu Val Val Val Pro Arg Asp Gln Thr
 515 520 525
 Pro Asp Glu Asn Asp Gln Val Val Val Lys Ile Thr Gly His Phe Tyr
 530 535 540
 Ala Cys Gln Val Ala Gln Arg Lys Ile Gln Glu Ile Leu Thr Gln Val
 545 550 555 560

Lys Gln His Gln Gln Gln Lys Ala Leu Gln Ser Gly Pro Pro Gln Ser
 565 570 575

Arg Arg Lys

<210> 177
 <211> 401
 <212> DNA
 <213> Homo sapiens

<400> 177
 atgccccgta aatgtcttca gtgttcttca gggtagttgg gatctcaaaa gatttggttc 60
 agatccaaac aaatacacat tctgtgtttt agctcagttg tttctaaaaa aagaaactgc 120
 cacacagcaa aaaattgttt actttgttgg acaaaccaaa tcagttctca aaaaatgacc 180
 ggtgcttata aaaagttata aatatcgagt agctctaaaa caaacacact gaccaagagg 240
 gaagtgaact tgtgcttagt atttacattg gatgccagtt ttgtaatcac tgacttatgt 300
 gcaaacctgg gcagaaattc tataaactct ttgctgtttt tgataacctg tttttgtttc 360
 attttgtttt gttttgtaaa aatgataaaa cttcagaaaa t 401

<210> 178
 <211> 561
 <212> DNA
 <213> Homo sapiens

<400> 178
 acgcctttca aggggtgtacg caaagcactc attgataccc ttttgatggg ctatgaaaca 60
 gcccgctatg ggacaggggt ctttggccag aatgagtacc tacgctatca ggaggccctg 120
 agtgagctgg ccactgcggt taaagcacga attgggagct ctcagcgaca tcaccagtca 180
 gcagccaaag acctaactca gtcccttgag gtctcccaaa caaccatcca ggtgacatac 240
 ctcccttcca gtcagaagag taaacgtgcc aagcaattcc ttgaattgaa gagctttaag 300
 gataactata acacattgga gagtactctg tgacggagct gaaggactct tgccgtagat 360
 taagccagtc agttgcaatg tgcaagacag gctgcttgcc gggccgcctt cggaacatct 420
 ggcccagcag gccagactg tatccatcca agttcccgtt gtatccagag ttcttagagc 480
 ttgtgtctaa agggtaatcc cccaaccctt ccttatgagc atttttagaa cattggctaa 540
 gactattttc cccagtagc g 561

<210> 179
 <211> 521
 <212> DNA
 <213> Homo sapiens

<400> 179
 cccaacgcgt ttgcaaatat tcccctggta gcctacttcc ttacccccga atattggtaa 60
 gatcgagcaa tggcttcagg acatgggttc tcttctcctg tgatcattca agtgctcact 120
 gcatgaagac tggcttgtct cagtgtttca acctcaccag ggctgtctct tgggtccacac 180
 ctgcctccct gttagtgcg tatgacagcc cccatcaaat gaccttggcc aagtcacggg 240
 ttctctgtgg tcaaggttgg ttggtgatt ggtggaaagt aggggtggacc aaaggaggcc 300
 acgtgagcag tcagcaccag ttctgcacca gcagcgcctc cgtcctagtg ggtgttcctg 360
 ttctccttgg ccttgggtgg gctagggcct gattcgggaa gatgcctttg caggaggggg 420
 aggataagt ggatctacca attgattctg gcaaaacaat ttctaagatt tttttgcttt 480
 atgtgggaaa cagatctaaa tctcatttta tgctgtattt t 521

<210> 180
 <211> 417
 <212> DNA
 <213> Homo sapiens

<400> 180
 ggtggaattc gccgaagatg gcggaggtgc aggtcctggt gcttgatggt cgaggccatc 60
 tcttgggccc cctggcgcc atcgtggcta aacagggtact gctgggcccg aaggtggtgg 120
 tcgtacgctg tgaaggcatc aacatttctg gcaatttcta cagaaacaag ttgaagtacc 180
 tggctttcct ccgcaagcgg atgaacacca acccttcccg aggcccctac cacttccggg 240
 cccccagccg catcttctgg cggaccgtgc gaggtatgct gcccacaaa accaagcgag 300
 gccagggcgc tctggaccgt ctcaagggtg ttgacggcat cccaccgcc tacgacaaga 360
 aaaagcggat ggtggttctt gctgccctca aggtcgtgcg tctgaagcct acaagaa 417

<210> 181
 <211> 283
 <212> DNA
 <213> Homo sapiens

<220>
 <221> unsure
 <222> (35)
 <223> n=A,T,C or G

<400> 181
 gattttcttct aaataggatg taaaacttct ttcanattac tcttctcag tcttgccctgc 60
 caagaactca agtgtaactg tgataaaata acctttccca ggtatattgg caggtatgtg 120
 tgtaatctca gaatacacag gtgacataga tatgatatga caactggtaa tgggtggattc 180
 atttacattg tttacacttc tatgaccagg ccttaaggga aggtcagttt tttaaaaaac 240
 caagtagtgt ctctctacct atctccagat acatgtcaaa aaa 283

<210> 182
 <211> 401
 <212> DNA
 <213> Homo sapiens

<400> 182
 atattcttgc tgcttatgca gctgacattg ttgccctccc taaagcaacc aagtagcctt 60
 tatttccac agtgaaagaa aacgctggcc tatcagttac attacaaaag gcagatttca 120
 agaggattga gtaagtagtt ggatggcttt cataaaaaaca agaattcaag aagaggattc 180
 atgctttaag aaacatttgt tatacatctc tcacaaatta tacctgggat aaaaactatg 240
 tagcaggcag tgtgttttcc ttccatgtct ctctgcacta cctgcagtgt gtctctctgag 300
 gctgcaagtc tgtcctatct gaattcccag cagaagcact aagaagctcc accctatcac 360
 ctagcagata aaactatggg gaaaacttaa atctgtgcat a 401

<210> 183
 <211> 366
 <212> DNA
 <213> Homo sapiens

<220>
 <221> unsure
 <222> (325)
 <223> n=A,T,C or G

<400> 183

```

accgtgtcca agtttttaga acccttggtta gccagaccga ggtgtcctgg tcaccgtttc 60
accatcatgc tttgatgttc ccctgtcttt ctctcttctg ctctcaagag caaagggttaa 120
tttaaggaca aagatgaagt cactgtaaac taatctgtca ttgtttttac cttccttttc 180
tttttcagtg cagaaattaa aagtaagtat aaagcaccgt gattgggagt gtttttgcgt 240
gtgtcggaat cactggtaaa tgttggctga gaacaatccc tccccttgca cttgtgaaaa 300
cactttgagc gctttaagag attancctga gaaataatta aatatctttt ctcttcaaaa 360
aaaaaa
366

```

<210> 184

<211> 370

<212> DNA

<213> Homo sapiens

<400> 184

```

tcttacttca aaagaaaaat aaacataaaa aataagttgc tggttcctaa caggaaaaat 60
tttaataatt gtactgagag aaactgctta cgtacacatt gcagatcaaa tatttgaggt 120
taaaatgtta gtctacatag atgggtgatt gtaactttat tgccattaaa agatttcaaa 180
ttgcattcat gcttctgtgt acacataatg aaaaatgggc aaataatgaa gatctctcct 240
tcagtctgct ctgtttaatt ctgctgtctg ctcttctcta atgctgcgtc cctaattgta 300
cacagtttag tgatatctag gagtataaag ttgtcgccca tcaataaaaa tcacaaagtt 360
ggtttaaaaa
370

```

<210> 185

<211> 107

<212> DNA

<213> Homo sapiens

<400> 185

```

ctcatattat tttccttttg agaaattgga aactctttct gttgctatta tattaataaa 60
gttggtgttt attttctggt agtcaccttc cccatttaaa aaaaaaa 107

```

<210> 186

<211> 309

<212> DNA

<213> Homo sapiens

<400> 186

```

gaaaggatgg ctctgggttg caccagagctg ggacttcatg ttcttctaga gagggccaca 60
agagggccac aggggtggcc gggagttgtc agctgatgcc tgctgagagg caggaattgt 120
gccagtgagt gacagtcatt agggagtgtc tcttcttggg gaggaagaa ggtagagcct 180
ttctgtctga atgaaaggcc aaggctacag tacagggccc cgccccagcc aggggtgttaa 240
tgcccacgta gtggaggcct ctggcagatc ctgcattcca aggtcactgg actgtacgtt 300
tttatggtt
309

```

<210> 187

<211> 477

<212> DNA

<213> Homo sapiens

<400> 187

```

ttcagtccta gcaagaagcg agaattctga gatcctccag aaagtcgagc agcaccacc 60
tccaacctcg ggccagtgtc ttcaggcttt actggggacc tgcgagctgg cctaattgtg 120

```

```

tggcctgcaa gccaggccat ccctggggcgc cacagacgag ctccgagcca ggtcaggctt 180
cggaggccac aagctcagcc tcaggcccag gcactgattg tggcagagg gccactaccc 240
aaggtctagc taggcccacg acctagttac ccagacagtg agaagcccct ggaaggcaga 300
aaagttggga gcatggcaga cagggaaggg aaacattttc agggaaaaga catgtatcac 360
atgtcttcag aagcaagtca ggtttcatgt aaccgagtg cctcttgctg gtccaaaagt 420
agcccagggc tgtagcacag gtttcacagt gattttgtgt tcagccgtga gtcacac 477

```

<210> 188

<211> 220

<212> DNA

<213> Homo sapiens

<400> 188

```

taaatatggt agatattaat attcctctta gatgaccagt gattccaatt gtcccaagtt 60
ttaaataagt accctgtgag tatgagataa attagtgaca atcagaacaa gtttcagtat 120
cagatgttca agaggaagtt gctattgcat tgattttaat atttgtacat aaacactgat 180
ttttttgagc attattttgt atttggttga ctttaatacc 220

```

<210> 189

<211> 417

<212> DNA

<213> Homo sapiens

<220>

<221> unsure

<222> (76)

<223> n=A,T,C or G

<221> unsure

<222> (77)

<223> n=A,T,C or G

<400> 189

```

accatcttga cagaggatac atgctcccaa aacgtttgtt accacactta aaaatcactg 60
ccatcattaa gcacnnttt caaaattata gccattcatg atttactttt tccagatgac 120
tatcattatt ctagtctttt gaatttgtaa ggggaaaaaa aacaaaaaca aaaacttacg 180
atgcactttt ctccagcaca tcagatttca aattgaaaat taaagacatg ctatggtaat 240
gcacttgcta gtactacaca ctttgtacaa caaaaaacag aggcaagaaa caacggaaag 300
agaaaagcct tcctttgttg gcccttaaac tgagtcaaga tctgaaatgt agagatgatc 360
tctgacgata cctgtatggt cttattgtgt aaataaaatt gctggtatga aatgaca 417

```

<210> 190

<211> 497

<212> DNA

<213> Homo sapiens

<400> 190

```

gcactgcggc gctctcccg cccgcgggtg ttgctgctgc tgccgctgct gctgggcctg 60
aacgcaggag ctgtcattga ctggcccaca gaggagggca aggaagtatg ggattatgtg 120
acggtccgca aggatgccta catgttcttg tggctctatt atgccaccaa ctctgcaag 180
aacttctcag aactgcccct ggtcatgtgg cttcaggggc gtccaggcgg ttctagcact 240
ggatttgga aacttgagga aattggggcc cttgacagtg atctcaaacc acggaaaacc 300
acctggctcc aggtgccag tctcctatct gtggataatc ccgtgggcac tgggttcagt 360
tatgtgaatg gtagtggtgc ctatgccaa gacctggcta tgggtggcttc agacatgatg 420
gttctcctga agaccttctt cagttgccac aaagaattcc agacagttcc attctacatt 480

```

ttctcagagt cctatgg

497

<210> 191

<211> 175

<212> DNA

<213> Homo sapiens

<400> 191

atgttgaata ttttgcttat taactttggt tattgtcttc tccctcgatt agaattattag 60
ctacttgagt acaaggattt gagcctgtta cattcactgc tgaatttttag gctcctggaa 120
gatacccagc attcaataga gaccacacaa taaatatatg tcaaataaaa aaaaa 175

<210> 192

<211> 526

<212> DNA

<213> Homo sapiens

<400> 192

agtaaacatt attatTTTTT ttatatTTTgc aaaggaaaca tatctaattcc ttcctataga 60
aagaacagta ttgctgtaat tccttttctt ttcttctca tttcctctgc ccttaaaag 120
attgaagaaa gagaaacttg tcaactcata tccacgttat ctagcaaagt acataagaat 180
ctatcactaa gtaatgtatc cttcagaatg tgttggttta ccagtgcacac cccatattca 240
tcacaaaatt aaagcaagaa gtccatagta atttatTTTgc taatagtgga tttttaatgc 300
tcagagtttc tgaggtcaaa ttttatcttt tcaattacaa gctctatgat cttaaataat 360
ttacttaatg tattttggtg tattttcttc aaattaatat tgggtgttcaa gactatatct 420
aattcctctg atcactttga gaaacaaact tttattaaat gtaaggcact tttctatgaa 480
ttttaaatat aaaaataaat attgttctga ttattactga aaaaaa 526

<210> 193

<211> 553

<212> DNA

<213> Homo sapiens

<220>

<221> unsure

<222> (290)

<223> n=A,T,C or G

<221> unsure

<222> (300)

<223> n=A,T,C or G

<221> unsure

<222> (411)

<223> n=A,T,C or G

<221> unsure

<222> (441)

<223> n=A,T,C or G

<400> 193

tccattgtgg tggaattcgc tctctggtaa aggcgtgcag gtgttggccg cggcctctga 60
gctgggatga gccgtgctcc cgggtggaagc aagggagccc agccggagcc atggccagta 120
cagtggtagc agttggactg accattgctg ctgcaggatt tgcaggccgt tacgttttgc 180
aagccatgaa gcatatggag cctcaagtaa aacaagtttt tcaaagccta ccaaaatctg 240
ccttcagtgg tggctattat agaggtgggt ttgaacccaa aatgacaaan cggaagcan 300
cattaatact aggtgtaagc cctactgcca ataaaggga aataagagat gctcatcgac 360


```

gaattatgct tttaaatcat cctgacaaag gaggatctcc ttatatagca nccaaaatca 420
atgaagctaa agatttacta naagggtcaag ctaaaaaatg aagtaaatgt atgatgaatt 480
ttaagttcgt attagtttat gtatatgagt actaagtttt tataataaaa tgcctcagag 540
ctacaatttt aaa 553

```

<210> 194

<211> 320

<212> DNA

<213> Homo sapiens

<400> 194

```

cccttcccaa tccatcagta aagaccccat ctgccttgtc catgccgttt cccaacaggg 60
atgtcacttg atatgagaat ctcaaactct aatgccttat aagcattcct tcctgtgtcc 120
attaagactc tgataattgt ctccccctca taggaatttc tcccaggaaa gaaatatatc 180
cccctctccg ttccatatca gaactaccgt ccccgatatt cccttcagag agattaaaga 240
ccagaaaaaa gtgagcctct tcatctgcac ctgtaatagt ttcagttcct attttcttcc 300
attgacccat atttatacct 320

```

<210> 195

<211> 320

<212> DNA

<213> Homo sapiens

<220>

<221> unsure

<222> (203)

<223> n=A,T,C or G

<221> unsure

<222> (218)

<223> n=A,T,C or G

<400> 195

```

aagcatgacc tggggaaatg gtcagacctt gtatttgttt tttggccttg aaagtagcaa 60
gtgaccagaa tctgccatgg caacaggctt taaaaaagac ccttaaaaag acactgtctc 120
aactgtggtg tttagcaccag ccagctctct gtacatttgc tagctttag ttttctaaga 180
ctgagtaaac ttcttatttt tanaaagggg aggctggntt gtaactttcc ttgtacttaa 240
ttgggtaaaa gtctttttcca caaaccacca tctattttgt gaactttgtt agtcactctt 300
tatttggtaa attatgaact 320

```

<210> 196

<211> 357

<212> DNA

<213> Homo sapiens

<220>

<221> unsure

<222> (36)

<223> n=A,T,C or G

<400> 196

```

atataaaaata atacgaaact ttaaaaagca ttggantgtc agtatgttga atcagtagtt 60
tcactttaac tgtaaacaaat ttcttaggac accatttggg ctagtttctg tgtaagtgtg 120
aatactacaa aaacttattt atactgttct tatgtcattt gttatattca tagatttata 180
tgatgatatg acatctggct aaaaagaaat tattgcaaaa ctaaccacta tgtacttttt 240

```

tataaatact gtatggacaa aaaatggcat tttttatatt aaattgttta gctctggcaa 300
 aaaaaaaaaa ttttaagagc tgggtactaat aaaggattat tatgactgtt aaaaaaa 357

<210> 197
 <211> 565
 <212> DNA
 <213> Homo sapiens

<220>
 <221> unsure
 <222> (27)
 <223> n=A,T,C or G

<400> 197
 tcagctgagt accatcagga tatttanccc ttttaagtgt gttttgggag tagaaaaacta 60
 aagcaacaat acttcctctt gacagctttg attggaatgg ggttattaga tcattcacct 120
 tggctctaca ctttttagga tgcttgggtga acataacacc acttataatg aacatccctg 180
 gttcctatat tttgggctat gtgggtagga attgttactt gttactgcag cagcagccct 240
 agaaaagtaag ccaggggctt cagatctaag ttagtcctaaa agctaaatga tttaaagtca 300
 agttgtaatg ctaggcataa gcactctata atacattaaa ttataggccg agcaattagg 360
 gaatgtttct gaaacattaa acttgatatt atgtcactaa aattctaaca caaacttaaa 420
 aaatgtgtct catacatatg ctgtactagg cttcatcatg catttctaaa tttgtgtatg 480
 atttgaatat atgaaagaat ttataacaaga gtgttattta aaattattaa aaataaatgt 540
 atataatttg tacctattgt aaaaa 565

<210> 198
 <211> 484
 <212> DNA
 <213> Homo sapiens

<400> 198
 tatgtaagta ttggtgtctg ctttaaaaaa ggagaccag acttcacctg tccttttttaa 60
 acatttgaga acagtgttac tctgagcagt tgggccacct tcaccttata cgacagctga 120
 ctgttggaatg tgtccattgt cgccagtttg gctgttgccc ggacaggaca ggacctccat 180
 tgggcgcagc agcaggtggc aggggtgttg cttgaggtgg gtggcagcgt ctggtcctcc 240
 tctctggtgc tttctgagag ggtctctaaa gcagagtgtg gttggcctgg gggaaggcag 300
 agcacgtatt tctccctct agtacctctg cattgtgag tgttccctct ggctttctga 360
 agggcagcag actcttgagt atactgcaga ggacatgctt tatcagtagg tcctgagggc 420
 tccaggggct caactgacca agtaacacag aagttggggt atgtggccta tttgggtcgg 480
 aaac 484

<210> 199
 <211> 429
 <212> DNA
 <213> Homo sapiens

<220>
 <221> unsure
 <222> (77)
 <223> n=A,T,C or G
 <221> unsure
 <222> (88)
 <223> n=A,T,C or G
 <221> unsure

<222> (134)
 <223> n=A,T,C or G
 <221> unsure
 <222> (151)
 <223> n=A,T,C or G
 <221> unsure
 <222> (189)
 <223> n=A,T,C or G
 <221> unsure
 <222> (227)
 <223> n=A,T,C or G
 <221> unsure
 <222> (274)
 <223> n=A,T,C or G
 <221> unsure
 <222> (319)
 <223> n=A,T,C or G

<400> 199
 gcttatgttt tttgttttaa cttttgtttt ttaacattta gaatattaca ttttgtatta 60
 tacagtacct ttctcanaca ttttgtanaa ttcatttcgg cagctcacta ggattttgct 120
 gaacattaaa aagngtgata gcgatattag ngccaatcaa atggaaaaaa ggtagtctta 180
 ataaacaana cacaacgttt ttataacaac tacttttaaa tattaanaaa actccttaat 240
 attgtttcct attaatgtatt attctttggg caanattttc tgatgctttt gattttctct 300
 caatttagca tttgctttng gtttttttct ctatttagca ttctgttaag gcacaaaaac 360
 tatgtactgt atgggaaatg ttgtaaatat taccttttcc acatttttaa cagacaactt 420
 tgaatccaa 429

<210> 200
 <211> 279
 <212> DNA
 <213> Homo sapiens

<400> 200
 gcttttttga ggaattacag ggaagctcct ggaattgtac atggatatct ttatccctag 60
 ggggaaatca aggagctggg caccctaat tctttatgga agtgtttaaa actattttta 120
 ttttattaca agtattacta gagtagtggt tctactctaa gatttcaaaa gtgcatttaa 180
 aatcatacat gttcccgctt gcaaatatat tgttattttg gtggagaaaa aaatagtata 240
 ttctacataa aaaattaaag atattaacta agaaaaaaa 279

<210> 201
 <211> 569
 <212> DNA
 <213> Homo sapiens

<400> 201
 taggtcagta ttttttagaaa ctcttaatag ctcatactct tgataccaaa agcagccctg 60
 attgttaaag cacacacctg cacaagaagc agtgatggtt gcatttacat ttcctgggtg 120
 cacaaaaaaa aattctcaaa aagcaaggac ttacgctttt tgcaaaagcct ttgagaagtt 180
 actggatcat aggaagctta taacaagaat ggaagattct taaataactc actttctttg 240
 gtatccagta acagttagtg ttcaaaaatg gtatgctgatt aataccagca ttgtgaacgc 300
 tgtacaacct tgtggttatt actaagcaag ttactactag cttctgaaaa gttagcttcat 360
 aattaatgtt atttatacac tgccttccat tacttttact ttgccctaag ctaatctcca 420
 aaatctgaaa tgctactcca atatcagaaa aaaaggggga ggtggaatta tatttctctg 480

gattttaaga gtacagagaa tcatgcacat ctctgattag ttcatatatg tctagtgtgt 540
aataaaagtc aaagatgaac tctcaaaaa 569

<210> 202
<211> 501
<212> DNA
<213> Homo sapiens

<400> 202
attaataggc ttaataattg ttggcaagga tccttttgct ttctttggca tgcaagctcc 60
tagcatctgg cagtggggcc aagaaaataa ggtttatgca tgtatgatgg ttttcttctt 120
gagcaacatg attgagaacc agtgtatgtc aacaggtgca tttgagataa ctttaaataga 180
tgtacctgtg tggctctaagc tgggaatctgg tcaccttcca tccatgcaac aacttgttca 240
aattcttgac aatgaaatga agctcaatgt gcatatggat tcaatcccac accatcgatc 300
atagcaccac ctatcagcac tgaaaactct tttgcattaa gggatcattg caagagcagc 360
gtgactgaca ttatgaaggc ctgtactgaa gacagcaagc tgtagtagaca gaccagatgc 420
tttcttggca ggctcgttgt acctcttggg aaacctcaat gcaagatagt gtttcagtgc 480
tggcatatct tgggaattctg c 501

<210> 203
<211> 261
<212> DNA
<213> Homo sapiens

<220>
<221> unsure
<222> (36)
<223> n=A,T,C or G
<221> unsure
<222> (96)
<223> n=A,T,C or G

<400> 203
gacaagctcc tggctcttgag atgtcttctc gttaangaga tgggcctttt ggaggtaaag 60
gataaaatga atgagttctg tcatgattca ctattntata acttgcatga cctttactgt 120
gttagctctt tgaatgttct tgaaatttta gactttcttt gtaaacaaat gatatgtcct 180
tatcattgta taaaagctgt tatgtgcaac agtgtggaga ttccttgtct gatttaataa 240
aataacttaa cactgaaaaa a 261

<210> 204
<211> 421
<212> DNA
<213> Homo sapiens

<400> 204
agcatctttt ctacaacggt aaaattgcag aagtagctta tcattaataaa acaacaacaa 60
caacaataac aataaatcct aagtgtaaat cagttattct accccctacc aaggatatca 120
gctgtttttt tccctttttt ctcttgggaa taattgtggg cttcttccca aatttctaca 180
gctcttttcc tcttctcatg cttgagcttc cctgtttgca cgcattgcgtg tgcaggactg 240
gcttggtgtg ttggactcgg ctccaggtgg aagcatgctt tcccttggtta ctgttgagga 300
aactcaaac ttcaagccct aggtgtagcc attttgtcaa gtcacaaact gtatttttgt 360
actggcatta acaaaaaaag aagataaaat attgtacat taaactttta taaaacttta 420
a 421

<210> 205
 <211> 460
 <212> DNA
 <213> Homo sapiens

<400> 205
 tactctcaca atgaaggacc tggaatgaaa aatctgtgtc taaacaagtc ctcttttagat 60
 tttagtgcaa atccagagcc agcgtcgggt gcctcgagta attctttcat gggtagcttt 120
 ggaaaagctc tcaggagacc tcacctagat gcctattcaa gctttggaca gccatcagat 180
 tgtcagccaa gagcctttta tttgaaagct cattcttccc cagacttgga ctctgggtca 240
 gaggaagatg ggaaagaaaag gacagatttt caggaagaaa atcacatttg tacctttaaa 300
 cagacttttag aaaactacag gactccaaat tttcagtcct atgacttgga cacatagact 360
 gaatgagacc aaaggaaaag cttaacatac tacctcaagg tgaactttta tttaaaagag 420
 agagaatctt atgtttttta aatggaggtta tgaattttta 460

<210> 206
 <211> 481
 <212> DNA
 <213> Homo sapiens

<400> 206
 tgtggtggaa ttcgggacgc cccagacccc tgactttttc ctgcgtgggc cgtctcctcc 60
 tgcggaagca gtgacctctg acccctggtg accttgcctt tgagtgcctt ttgaacgctg 120
 gtcccgcggg acttggtttt ctcaagctct gtctgtccaa agacgtccg gtcgaggtcc 180
 cgctgtccct ggggtggatac ttgaacccca gaagccctc tgtgtgtctg tgtccggagg 240
 cggccttccc atctgcctgc ccaccggag ctctttccgc cggcgcaggg tcccaagccc 300
 acctcccgcc ctcatgctct cggtgtgctg ctgggcacgt cctgcacaca caatgcaagt 360
 cctggcctcc gcgcccgcgc gccacgcga gccgtaccgc ccgccaactc tgttatttat 420
 ggtgtgaccc cctggagggtg ccctcgcccc accggggcta tttattgttt aatttatttg 480
 t 481

<210> 207
 <211> 605
 <212> DNA
 <213> Homo sapiens

<400> 207
 accttttttg gattcagggc tctcacaat taaaatgagt gtaatgaaac aaggtgaaaa 60
 tatagaagca tccctttgta tactgttttg ctacttacag tgtacttggc attgctttat 120
 ctactggat tctcacggta ggattttctga gatcttaatc taagctccaa agttgtctac 180
 ttttttgatc ctagggtgct ccttttgttt tacagagcag ggtcacttga tttgctagct 240
 ggtggcagaa ttggcaccat taaccaggtc tgactgacca ccagtcagag gcactttatt 300
 tgtatcatga aatgatttga aatcatttga aagcagcgaa gtctgataat gaatgccagc 360
 tttccttggt ctttgataac aaagactcca aatattctgg agaacttga taaaagtttg 420
 aagggctaga ttgggatttg aagacaaaat tgtaggaaat cttacatttt tgcaataaca 480
 aacattaatg aaagcaaaac attataaaaag taattttaat tcaccacata cttatcaatt 540
 tcttgatgct tccaaatgac atctaccaga tatggttttg tggacatctt tttctgttta 600
 cataa 605

<210> 208
 <211> 655
 <212> DNA
 <213> Homo sapiens

```

<400> 208
ggcgttggttc tggattcccc tcgtaactta aagggaact ttcacaatgt ccggagccct 60
tgatgtcctg caaatgaagg aggaggatgt ccttaagttc cttgcagcag gaaccactt 120
aggtggcacc aatcttgact tccagatgga acagtacatc tataaaagga aaagtgatgg 180
catctatata ataaatctca agaggacctg ggagaagctt ctgctggcag ctgctgcaat 240
tgttgccatt gaaaaccctg ctgatgtcag tgttatatcc tccaggaata ctggccagag 300
ggctgtgctg aagtttgctg ctgccactgg agccactcca attgctggcc gcttcaactc 360
tggaaccttc actaaccaga tccaggcagc cttccgggag ccacggcttc ttgtggttac 420
tgaccccagg gctgaccacc agcctctcac ggaggcatct tatgttaacc tacctaccat 480
tgcgctgtgt aacacagatt ctctctcgcg ctatgtggac attgccatcc catgcaacaa 540
caagggagct cactcagtg gtttgatgtg gtggatgctg gctcgggaag ttctgcgcac 600
gcgtggcacc atttcccggtg aacacccatg ggaggtcacg cctgatctgt acttc 655

```

<210> 209

<211> 621

<212> DNA

<213> Homo sapiens

```

<400> 209
catttagaac atggttatca tccaagacta ctctacctg caacattgaa ctcccaagag 60
caaatccaca ttcctcttga gttctgcagc ttctgtgtaa atagggcagc tgtcgtctat 120
gccgtagaat cacatgatct gaggaccatt catggaagct gctaaatagc ctagtctggg 180
gagtcttcca taaagttttg catggagcaa acaaacagga ttaaactagg ttgtgttcct 240
tcagccctct aaaagcatag ggcttagcct gcaggcttcc ttgggctttc tctgtgtgtg 300
tagttttgta aacactatag catctgttaa gatccagtgt ccatggaaac cttcccat 360
gccgtgactc tggactatat cagtttttgg aaagcagggt tcctctgcct gctaacaagc 420
ccacgtggac cagtctgaat gtctttcctt tacacctatg tttttaata gtcaaacttc 480
aagaaacaat ctaaacaagt ttctgttgca tatgtgtttg tgaacttgta ttgtattta 540
gtaggcttct atattgcatt taacttgttt ttgtaactcc tgattcttcc ttttcggata 600
ctattgatga ataaagaaat t
621

```

<210> 210

<211> 533

<212> DNA

<213> Homo sapiens

<220>

<221> unsure

<222> (20)

<223> n=A,T,C or G

<221> unsure

<222> (21)

<223> n=A,T,C or G

<221> unsure

<222> (61)

<223> n=A,T,C or G

<400> 210

```

cgcttgggg agccggcggn ngagtccggg acgtggagac ccgggggtccc ggcagccggg 60
nggcccgcg gccagggtg gggatgcacc gccgcgggt gggagctggc gccatcgcca 120
agaagaaact tgcagaggcc aagtataagg agcgaggac ggtcttggct gaggaccagc 180
tagcccagat gtcaaagcag ttggacatgt tcaagaccaa cctggaggaa ttgtccagca 240
aacacaagca ggagatccgg aagaatcctg agttccgtgt gcagttccag gacatgtgtg 300
caaccattgg cgtggatccg ctggcctctg gaaaaggatt ttggtctgag atgctgggcg 360

```

```

tgggggactt ctattacgaa ctaggtgtcc aaattatcga agtgtgcctg gcgctgaagc 420
atcggaatgg aggtctgata actttggagg aactacatca acaggtgttg aagggaaggg 480
gcaagttcgc ccaggatgtc agtcaagatg acctgatcag agccatcaag aaa 533

```

```

<210> 211
<211> 451
<212> DNA
<213> Homo sapiens

```

```

<400> 211
ttagcttgag ccgagaacga ggcgagaaag ctggagaccg aggagaccgc ctagagcgga 60
gtgaacgggg aggggaccgt ggggaccggc ttgatcgtgc gcggacacct gctaccaagc 120
ggagcttcag caaggaagtg gaggagcggg gtagagaacg gccctcccag cctgaggggc 180
tgcgcaaggc agctagcctc acggaggatc gggaccgtgg gcgggatgcc gtgaagcgag 240
aagctgccct acccccagtg agccccctga aggcggctct ctctgaggag gagttagaga 300
agaaatccaa ggctatcatt gaggaatata tccatctcaa tgacatgaaa gaggcagtc 360
agtgcggtgca ggagctggcc tcaccctcct tgctcttcat ctttgtacgg catggtgtcg 420
agtctacgct ggagcgcagt gccattgctc g 451

```

```

<210> 212
<211> 471
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> unsure
<222> (54)
<223> n=A,T,C or G

```

```

<400> 212
gtgattattc ttgatcaggg agaagatcat ttagatttgt tttgcattcc ttanaatgga 60
gggcaacatt ccacagctgc cctggctgtg atgagtgtcc ttgcaggggc cggagtagga 120
gcaactgggt gggggcggaa ttgggggttac tcgatgtaag ggattccttg ttgttgtgtt 180
gagatccagt gcagttgtga tttctgtgga tcccagcttg gttccaggaa ttttgtgtga 240
ttggcttaaa tccagttttc aatcttcgac agctgggctg gaacgtgaac tcagtagctg 300
aacctgtctg acccggtcac gttcttgat cctcagaact ctttgctctt gtcggggtgg 360
gggtgggaac tcacgtgggg agcgggtggc gagaaaatgt aaggattctg gaatacatat 420
tccatgggac tttccttccc tctcctgctt cctcttttcc tgctccctaa c 471

```

```

<210> 213
<211> 511
<212> DNA
<213> Homo sapiens

```

```

<220>
<221> unsure
<222> (27)
<223> n=A,T,C or G
<221> unsure
<222> (63)
<223> n=A,T,C or G
<221> unsure
<222> (337)
<223> n=A,T,C or G

```

<221> unsure
 <222> (442)
 <223> n=A,T,C or G

<400> 213
 ctaattagaa acttgctgta cttttntttt tcttttaggg gtcaaggacc ctctttatag 60
 ctncatttg cctacaataa attattgcag cagtttgcaa tactaaaata ttttttatag 120
 actttatatt tttccttttg ataaagggat gctgcatagt agagttagtg taattaaact 180
 atctcagccg tttccctgct ttcccttctg ctccatagtc ctcatgtcc ttccagggag 240
 ctcttttaat cttaaagttc tacatttcat gctcttagtc aaattctgtt acctttttaa 300
 taactcttcc cactgcatat ttccatcttg aattggnggt tctaaattct gaaactgtag 360
 ttgagataca gctattttaat atttctggga gatgtgcac cctcttcttt gtggtgccc 420
 aaggttgttt tgcgtaactg anactccttg atatgcttca gagaatttag gcaaactg 480
 gccatggccg tgggagtact gggagtaaaa t 511

<210> 214
 <211> 521
 <212> DNA
 <213> Homo sapiens

<400> 214
 agcattgcc aataatccct aattttccac taaaaatata atgaaatgat gttaagcttt 60
 ttgaaaagtt taggttaaac ctactgttgt tagattaatg tatttgttg ttccctttat 120
 ctggaatgtg gcattagctt ttttatttta accctcttta attcttattc aattccatga 180
 ctttaaggtg gagagctaaa cactgggatt tttggataac agactgacag ttttgcataa 240
 ttataatcgg cattgtacat agaaaggata tggctacctt ttgttaaadc tgcactttct 300
 aaatatcaaa aaagggaaat gaagtataaa tcaatttttg tataatctgt ttgaaacatg 360
 agttttatct gcttaatat agggctttgc ccttttctg taagtctctt gggatcctgt 420
 gtagaagctg ttctcattaa acaccaaaca gttaagtcca ttctctggta ctactacaa 480
 attcggtttc atattctact taacaattta aataaactga a 521

<210> 215
 <211> 381
 <212> DNA
 <213> Homo sapiens

<220>
 <221> unsure
 <222> (17)
 <223> n=A,T,C or G
 <221> unsure
 <222> (20)
 <223> n=A,T,C or G
 <221> unsure
 <222> (60)
 <223> n=A,T,C or G
 <221> unsure
 <222> (61)
 <223> n=A,T,C or G
 <221> unsure
 <222> (365)
 <223> n=A,T,C or G

<400> 215


```

gagcggagag cggaccngtn agagccctga gcagcccccac cggcggccgcc ggcctagttn 60
ncatcacacc ccgggaggag ccgcagctgc cgcagccggc ccagtcacc atcacccgaa 120
ccatgagcag cgaggccgag acccagcagc cggcggccgcc cggcggccgcc gggcggccgcc 180
tcagcggcgc cgacaccaag cccggcacta cgggcagcgg cgcagggagc ggtggcccg 240
gcggcctcac atcggcggcg cctgccggcg gggacaagaa ggtcatcgca acgaagggtt 300
tggaacagt aaaatggttc aatgtaagga acggatatgg ttcatcaac aggaatgaca 360
ccaangaaga tgtatttgta c                                     381

```

<210> 216

<211> 425

<212> DNA

<213> Homo sapiens

<400> 216

```

ttactaacta ggtcattcaa ggaagtcaag ttaacttaaa catgtcacct aaatgcactt 60
gatgggtgtg aaatgtccac cttcttaaat ttttaagatg aacttagttc taaagaagat 120
aacaggccaa tcttgaaggt actccctgtt tgctgcagaa tgtcagatat tttggatgtt 180
gcataagagt cctatttgcc ccagttaatt caacttttgt ctgcctgttt tgtggactgg 240
ctggctctgt tagaactctg tccaaaaagt gcatggaata taacttgtaa agcttccac 300
aattgacaat atatatgcat gtgttttaaac caaatccaga aagcttaaac aatagagctg 360
cataatagta tttattaaag aatcacaact gtaaacaatga gaataactta aggattctag 420
tttag                                             425

```

<210> 217

<211> 181

<212> DNA

<213> Homo sapiens

<400> 217

```

gagaaaccaa atgatagggt gtagagcctg atgactccaa acaaagccat caccgcatt 60
cttcctcctt cttctggtgc tacagctcca agggcccttc accttcatgt ctgaaatgga 120
actttggcct tttcagtgga agaatatgtt gaaggtttca ttttgttcta gaaaaaaaaa 180
a                                             181

```

<210> 218

<211> 405

<212> DNA

<213> Homo sapiens

<400> 218

```

caggccttcc agttcactga caaacatggg gaagtgtgcc cagctggctg gaaacctggc 60
agtgatacca tcaagcctga tgtccaaaag agcaaagaat atttctccaa gcagaagtga 120
gcgtggggt gtttttagtg caggctgcgg tgggcagcca tgagaacaaa acctcttctg 180
tattttttt ttccattagt aaaacacaag acttcagatt cagccgaatt gtggtgtctt 240
acaaggcagg ctttctctac agggggtgga gagaccagc tttcttctt tggtaggaat 300
ggcctgagtt ggcgttggtg gcaggctact ggtttgatg atgtattagt agagcaacc 360
attaatctt ttagtattgt attaaacttg aactgagaaa aaaaaa                                     405

```

<210> 219

<211> 216

<212> DNA

<213> Homo sapiens

<220>

<221> unsure
 <222> (207)
 <223> n=A,T,C or G
 <221> unsure
 <222> (210)
 <223> n=A,T,C or G

<400> 219
 actccaagag ttagggcagc agagtggagc gatttagaaa gaacatttta aaacaatcag 60
 ttaattttacc atgtaaaatt gctgtaaatg ataatgtgta cagattttct gttcaaatat 120
 tcaattgtaa acttcttggt aagactgtta cgtttctatt gcttttgtat gggatattgc 180
 aaaaataaaa aggaaagaac cctcttnaan aaaaaa 216

<210> 220
 <211> 380
 <212> DNA
 <213> Homo sapiens

<400> 220
 cttacaaatt gcccccatgt gtaggggaca cagaaccctt tgagaaaact tagatttttg 60
 tctgtacaaa gtctttgcct ttttccttct tcattttttt ccagtacatt aaatttgtca 120
 atttcactct tgagggaaac tgattagatg ggttggtgtt gtgttctgat ggagaaaaca 180
 gcacccaag gactcagaag atgattttta cagttcagaa cagatgtgtg caatattggt 240
 gcatgtaata atgttgagtg gcagtcaaaa gtcattgatt ttatcttagt tcttcattac 300
 tgcattgaaa aggaaaacct gtctgagaaa atgcctgaca gtttaattta aaactatggt 360
 gtaagtcttt gacaaaaaaa 380

<210> 221
 <211> 398
 <212> DNA
 <213> Homo sapiens

<400> 221
 ggttagtaag ctgtcgactt tgtaaaaaag ttaaaaatga aaaaaaaagg aaaaatgaat 60
 tgtatattta atgaatgaac atgtacaatt tgccactggg aggaggttcc tttttgttgg 120
 gtgagctctg aagtgaattt cactgatgtt gatattcatt gtgtgtagtt ttatttcggt 180
 cccagcccg tttcctttta ttttgagct aatgccagct gcgtgtctag ttttgagtgc 240
 agtaaaatag aatcagcaaa tcaactcttat ttttcactct tttccggtat tttttgggtt 300
 gtttctgtgg gagcagtgt caccaactct tctgtatat tgcctttttg ctggaaaatg 360
 ttgtatgttg aataaaattt tctataaaaa ttaaaaaa 398

<210> 222
 <211> 301
 <212> DNA
 <213> Homo sapiens

<220>
 <221> unsure
 <222> (49)
 <223> n=A,T,C or G
 <221> unsure
 <222> (64)
 <223> n=A,T,C or G

<400> 222

```

ttcgataatt gatctcatgg gctttccctg gaggaaggt tttttttgnt gtttattttt 60
taanaacttg aaacttgtaa actgagatgt ctgtagcttt ttgcccac tgtagtgat 120
gtgaagattt caaaacctga gagcactttt tctttgttta gaattatgag aaaggcacta 180
gatgacttta ggatttgcac ttttcccttt attgctcat ttcttgtagc gccttggtgg 240
ggagggaat ctgtttattt tttcctacaa ataaaaagct aagattctat atcgcaaaaa 300
a 301

```

<210> 223

<211> 200

<212> DNA

<213> Homo sapiens

<400> 223

```

gtaagtgtt aggaagaaac tttgcaaaca tttaatgagg atacactgtt catttttaaa 60
attccttcac actgtaattt aatgtgtttt atattctttt gtagtaaaac aacataactc 120
agatttctac aggagacagt ggttttattt ggattgtctt ctgtaatagg tttcaataaa 180
gctggatgaa cttaaaaaaa 200

```

<210> 224

<211> 385

<212> DNA

<213> Homo sapiens

<400> 224

```

gaaaggtttg atccggactc aaagaaagca aaggagtgtg agccgccatc tgctggagca 60
gtgttaactg caagacctgg acaagagatt cgtcagcgaa ctgcagctca aagaaacctt 120
tctccaacac cagcaagccc taaccagggc cctcctccac aagtccagt atctcctgga 180
ccaccaaaagg acagtctctg ccttggtgga cccccagaaa ggactgttac tccagcccta 240
tcatcaaatg tgttaccaag acatcttgga tcccctgcta ctccagtgcc tggaatgggt 300
aaacagagca cttaatgtta tttacagttt atattgtttt ctctgggttac caataaaacg 360
ggccattttc aggtggttaa aaaaa 385

```

<210> 225

<211> 560

<212> PRT

<213> Homo sapien

<400> 225

```

Met Glu Cys Leu Tyr Tyr Phe Leu Gly Phe Leu Leu Leu Ala Ala Arg
 1          5          10          15
Leu Pro Leu Asp Ala Ala Lys Arg Phe His Asp Val Leu Gly Asn Glu
 20          25          30
Arg Pro Ser Ala Tyr Met Arg Glu His Asn Gln Leu Asn Gly Trp Ser
 35          40          45
Ser Asp Glu Asn Asp Trp Asn Glu Lys Leu Tyr Pro Val Trp Lys Arg
 50          55          60
Gly Asp Met Arg Trp Lys Asn Ser Trp Lys Gly Arg Val Gln Ala
 65          70          75          80
Val Leu Thr Ser Asp Ser Pro Ala Leu Val Gly Ser Asn Ile Thr Phe
          85          90          95
Ala Val Asn Leu Ile Phe Pro Arg Cys Gln Lys Glu Asp Ala Asn Gly
100          105          110

```

Asn Ile Val Tyr Glu Lys Asn Cys Arg Asn Glu Ala Gly Leu Ser Ala
 115 120 125
 Asp Pro Tyr Val Tyr Asn Trp Thr Ala Trp Ser Glu Asp Ser Asp Gly
 130 135 140
 Glu Asn Gly Thr Gly Gln Ser His His Asn Val Phe Pro Asp Gly Lys
 145 150 155 160
 Pro Phe Pro His His Pro Gly Trp Arg Arg Trp Asn Phe Ile Tyr Val
 165 170 175
 Phe His Thr Leu Gly Gln Tyr Phe Gln Lys Leu Gly Arg Cys Ser Val
 180 185 190
 Arg Val Ser Val Asn Thr Ala Asn Val Thr Leu Gly Pro Gln Leu Met
 195 200 205
 Glu Val Thr Val Tyr Arg Arg His Gly Arg Ala Tyr Val Pro Ile Ala
 210 215 220
 Gln Val Lys Asp Val Tyr Val Val Thr Asp Gln Ile Pro Val Phe Val
 225 230 235 240
 Thr Met Phe Gln Lys Asn Asp Arg Asn Ser Ser Asp Glu Thr Phe Leu
 245 250 255
 Lys Asp Leu Pro Ile Met Phe Asp Val Leu Ile His Asp Pro Ser His
 260 265 270
 Phe Leu Asn Tyr Ser Thr Ile Asn Tyr Lys Trp Ser Phe Gly Asp Asn
 275 280 285
 Thr Gly Leu Phe Val Ser Thr Asn His Thr Val Asn His Thr Tyr Val
 290 295 300
 Leu Asn Gly Thr Phe Ser Leu Asn Leu Thr Val Lys Ala Ala Ala Pro
 305 310 315 320
 Gly Pro Cys Pro Pro Pro Pro Pro Pro Arg Pro Ser Lys Pro Thr
 325 330 335
 Pro Ser Leu Gly Pro Ala Gly Asp Asn Pro Leu Glu Leu Ser Arg Ile
 340 345 350
 Pro Asp Glu Asn Cys Gln Ile Asn Arg Tyr Gly His Phe Gln Ala Thr
 355 360 365
 Ile Thr Ile Val Glu Gly Ile Leu Glu Val Asn Ile Ile Gln Met Thr
 370 375 380
 Asp Val Leu Met Pro Val Pro Trp Pro Glu Ser Ser Leu Ile Asp Phe
 385 390 395 400
 Val Val Thr Cys Gln Gly Ser Ile Pro Thr Glu Val Cys Thr Ile Ile
 405 410 415
 Ser Asp Pro Thr Cys Glu Ile Thr Gln Asn Thr Val Cys Ser Pro Val
 420 425 430
 Asp Val Asp Glu Met Cys Leu Leu Thr Val Arg Arg Thr Phe Asn Gly
 435 440 445
 Ser Gly Thr Tyr Cys Val Asn Leu Thr Leu Gly Asp Asp Thr Ser Leu
 450 455 460
 Ala Leu Thr Ser Thr Leu Ile Ser Val Pro Asp Arg Asp Pro Ala Ser
 465 470 475 480
 Pro Leu Arg Met Ala Asn Ser Ala Leu Ile Ser Val Gly Cys Leu Ala
 485 490 495
 Ile Phe Val Thr Val Ile Ser Leu Leu Val Tyr Lys Lys His Lys Glu
 500 505 510
 Tyr Asn Pro Ile Glu Asn Ser Pro Gly Asn Val Val Arg Ser Lys Gly
 515 520 525
 Leu Ser Val Phe Leu Asn Arg Ala Lys Ala Val Phe Phe Pro Gly Asn
 530 535 540

Gln Glu Lys Asp Pro Leu Leu Lys Asn Gln Glu Phe Lys Gly Val Ser
 545 550 555 560

<210> 226
 <211> 9
 <212> PRT
 <213> Homo sapien

<400> 226
 Ile Leu Ile Pro Ala Thr Trp Lys Ala
 1 5

<210> 227
 <211> 9
 <212> PRT
 <213> Homo sapien

<400> 227
 Phe Leu Leu Asn Asp Asn Leu Thr Ala
 1 5

<210> 228
 <211> 9
 <212> PRT
 <213> Homo sapien

<400> 228
 Leu Leu Gly Asn Cys Leu Pro Thr Val
 1 5

<210> 229
 <211> 10
 <212> PRT
 <213> Homo sapien

<400> 229
 Lys Leu Leu Gly Asn Cys Leu Pro Thr Val
 1 5 10

<210> 230
 <211> 10
 <212> PRT
 <213> Homo sapien

<400> 230
 Arg Leu Thr Gly Gly Leu Lys Phe Phe Val
 1 5 10

<210> 231
 <211> 9
 <212> PRT
 <213> Homo sapien

<400> 231

Ser Leu Gln Ala Leu Lys Val Thr Val
1 5

<210> 232

<211> 20

<212> PRT

<213> Homo sapiens

<400> 232

Ala Gly Ala Asp Val Ile Lys Asn Asp Gly Ile Tyr Ser Arg Tyr Phe
5 10 15

Phe Ser Phe Ala
20

<210> 233

<211> 21

<212> PRT

<213> Homo sapiens

<400> 233

Phe Phe Ser Phe Ala Ala Asn Gly Arg Tyr Ser Leu Lys Val His Val
5 10 15

Asn His Ser Pro Ser
20

<210> 234

<211> 20

<212> PRT

<213> Homo sapiens

<400> 234

Phe Leu Val Thr Trp Gln Ala Ser Gly Pro Pro Glu Ile Ile Leu Phe
5 10 15

Asp Pro Asp Gly
20

<210> 235

<211> 20

<212> PRT

<213> Homo sapiens

<400> 235

Leu Gln Ser Ala Val Ser Asn Ile Ala Gln Ala Pro Leu Phe Ile Pro
5 10 15

Pro Asn Ser Asp
20

<210> 236
 <211> 20
 <212> PRT
 <213> Homo sapiens

<400> 236
 Ile Gln Asp Asp Phe Asn Asn Ala Ile Leu Val Asn Thr Ser Lys Arg
 5 10 15
 Asn Pro Gln Gln
 20

<210> 237
 <211> 21
 <212> PRT
 <213> Homo sapiens

<400> 237
 Arg Asn Ser Leu Gln Ser Ala Val Ser Asn Ile Ala Gln Ala Pro Leu
 5 10 15
 Phe Ile Pro Pro Asn
 20

<210> 238
 <211> 20
 <212> PRT
 <213> Homo sapiens

<400> 238
 Thr His Glu Ser His Arg Ile Tyr Val Ala Ile Arg Ala Met Asp Arg
 5 10 15
 Asn Ser Leu Gln
 20

<210> 239
 <211> 20
 <212> PRT
 <213> Homo sapiens

<400> 239
 Arg Asn Pro Gln Gln Ala Gly Ile Arg Glu Ile Phe Thr Phe Ser Pro
 5 10 15
 Gln Ile Ser Thr
 20

<210> 240
 <211> 21
 <212> PRT
 <213> Homo sapiens

<400> 240
 Gly Gln Ala Thr Ser Tyr Glu Ile Arg Met Ser Lys Ser Leu Gln Asn
 5 10 15
 Ile Gln Asp Asp Phe
 20

<210> 241
 <211> 20
 <212> PRT
 <213> Homo sapiens

<400> 241
 Glu Arg Lys Trp Gly Phe Ser Arg Val Ser Ser Gly Gly Ser Phe Ser
 5 10 15
 Val Leu Gly Val
 20

<210> 242
 <211> 20
 <212> PRT
 <213> Homo sapiens

<400> 242
 Gly Ser His Ala Met Tyr Val Pro Gly Tyr Thr Ala Asn Gly Asn Ile
 5 10 15
 Gln Met Asn Ala
 20

<210> 243
 <211> 20
 <212> PRT
 <213> Homo sapiens

<400> 243
 Val Asn His Ser Pro Ser Ile Ser Thr Pro Ala His Ser Ile Pro Gly
 5 10 15
 Ser His Ala Met
 20

<210> 244
 <211> 20

<212> PRT

<213> Homo sapiens

<400> 244

Ala Val Pro Pro Ala Thr Val Glu Ala Phe Val Glu Arg Asp Ser Leu
 5 10 15

His Phe Pro His
 20

<210> 245

<211> 20

<212> PRT

<213> Homo sapiens

<400> 245

Lys Pro Gly His Trp Thr Tyr Thr Leu Asn Asn Thr His His Ser Leu
 5 10 15

Gln Ala Leu Lys
 20

<210> 246

<211> 20

<212> PRT

<213> Homo sapiens

<400> 246

Asn Leu Thr Phe Arg Thr Ala Ser Leu Trp Ile Pro Gly Thr Ala Lys
 5 10 15

Pro Gly His Trp
 20

<210> 247

<211> 20

<212> PRT

<213> Homo sapiens

<400> 247

Leu His Phe Pro His Pro Val Met Ile Tyr Ala Asn Val Lys Gln Gly
 5 10 15

Phe Tyr Pro Ile
 20

<210> 248

<211> 20

<212> PRT

<213> Homo sapiens

<400> 248

Pro Glu Thr Gly Asp Pro Val Thr Leu Arg Leu Leu Asp Asp Gly Ala
 5 10 15

Gly Ala Asp Val
 20

<210> 249

<211> 20

<212> PRT

<213> Homo sapiens

<400> 249

Gly Phe Tyr Pro Ile Leu Asn Ala Thr Val Thr Ala Thr Val Glu Pro
 5 10 15

Glu Thr Gly Asp
 20

<210> 250

<211> 20

<212> PRT

<213> Homo sapiens

<400> 250

Phe Asp Pro Asp Gly Arg Lys Tyr Tyr Thr Asn Asn Phe Ile Thr Asn
 5 10 15

Leu Thr Phe Arg
 20

<210> 251

<211> 20

<212> PRT

<213> Homo sapiens

<400> 251

Leu Gln Ala Leu Lys Val Thr Val Thr Ser Arg Ala Ser Asn Ser Ala
 5 10 15

Val Pro Pro Ala
 20

<210> 252

<211> 153

<212> PRT

<213> Homo sapien

<400> 252

Met Ala Ser Val Arg Val Ala Ala Tyr Phe Glu Asn Phe Leu Ala Ala
 1 5 10 15
 Trp Arg Pro Val Lys Ala Ser Asp Gly Asp Tyr Tyr Thr Leu Ala Val
 20 25 30
 Pro Met Gly Asp Val Pro Met Asp Gly Ile Ser Val Ala Asp Ile Gly
 35 40 45
 Ala Ala Val Ser Ser Ile Phe Asn Ser Pro Glu Glu Phe Leu Gly Lys
 50 55 60
 Ala Val Gly Leu Ser Ala Glu Ala Leu Thr Ile Gln Gln Tyr Ala Asp
 65 70 75 80
 Val Leu Ser Lys Ala Leu Gly Lys Glu Val Arg Asp Ala Lys Ile Thr
 85 90 95
 Pro Glu Ala Phe Glu Lys Leu Gly Phe Pro Ala Ala Lys Glu Ile Ala
 100 105 110
 Asn Met Cys Arg Phe Tyr Glu Met Lys Pro Asp Arg Asp Val Asn Leu
 115 120 125
 Thr His Gln Leu Asn Pro Lys Val Lys Ser Phe Ser Gln Phe Ile Ser
 130 135 140
 Glu Asn Gln Gly Ala Phe Lys Gly Met
 145 150

<210> 253
 <211> 462
 <212> DNA
 <213> Homo sapien

<400> 253

atggccagtg	tccgcggtggc	ggcctacttt	gaaaactttc	tccgcgcggtg	gcgggcccggtg	60
aaagcctctg	atggagatta	ctacaccttg	gctgtaccga	tgggagatgt	accaatggat	120
ggtatctctg	ttgctgatat	tggagcagcc	gtctctagca	tttttaattc	tccagaggaa	180
tttttaggca	aggcgtggg	gctcagtgc	gaagcaactaa	caatacagca	atatgctgat	240
gttttgtcca	aggctttggg	gaaagaagtc	cgagatgcaa	agattacccc	ggaagctttc	300
gagaagctgg	gattccctgc	agcaaaggaa	atagccaata	tgtgtcgttt	ctatgaaatg	360
aagccagacc	gagatgtcaa	tctcacccac	caactaaatc	ccaaagtcaa	aagcttcagc	420
cagtttatct	cagagaacca	gggagccttc	aagggcattgt	ag		462

<210> 254
 <211> 8031
 <212> DNA
 <213> Homo sapien

<400> 254

tggcgaaatg	gacgcgccct	gtagcggcgc	attaagcgcg	gcggggtgtgg	tggttacgcg	60
cagcgtgacc	gctacacttg	ccagcgccct	agcgcccgt	cctttcgctt	tcttcccttc	120
ctttctcgcc	acgttcgccc	gctttccccg	tcaagctcta	aatcgggggc	tccctttagg	180
gttccgattt	agtgccttac	ggcacctcga	ccccaaaaaa	cttgattagg	gtgatggttc	240
acgtagtggg	ccatcgccct	gatagacggg	ttttcgccct	ttgacgttgg	agtccacggt	300
ctttaatagt	ggactccttg	tccaaactgg	aacaacactc	aaccctatct	cggtctattc	360
ttttgattta	taagggattt	tgccgatttc	ggcctatttg	ttaaaaaatg	agctgattta	420
acaaaaattt	aacgcgaatt	ttaacaaaa	attaacgttt	acaatttcag	gtggcacttt	480
tcggggaaat	gtgcgcggaa	cccctatttg	tttatttttc	ttaatacatt	caaatatgta	540
tccgctcatg	aattaattct	tagaaaaact	catcgagcat	caaatgaaac	tgcaatttat	600
tcatatcagg	attatcaata	ccatatTTTT	gaaaaagccg	tttctgtaat	gaaggagaaa	660
actcaccgag	gcagttccat	aggatggcaa	gacccctgga	tccgtctgcg	attccgactc	720

gtccaacatc	aatacaacot	attaatttcc	cctcgtcaaa	aataaggtta	tcaagtgaga	780
aatcaccatg	agtgcagact	gaatccggtg	agaattggcaa	aagtttatgc	atttttttcc	840
agacttggtc	aacaggccag	ccattacgct	cgtcatcaaa	atcactcgca	tcaaccaaac	900
cgttattcat	tcgtgattgc	gcctgagcga	gacgaaatac	gcgatcgctg	ttaaaaggac	960
aattacaaac	aggaatcgaa	tgcaaccggc	gcaggaacac	tgccagcgca	tcaacaatat	1020
tttcacctga	atcaggatat	tcttctaata	cctggaatgc	tgttttcccg	gggatcgag	1080
tggtgagtaa	ccatgcatca	tcaggagtac	ggataaaatg	cttgatggtc	ggaagaggca	1140
taaattccgt	cagccagttt	agtctgacca	tctcatctgt	aacatcattg	gcaacgctac	1200
ctttgccatg	tttcagaaac	aactctggcg	catcgggctt	cccatacaat	cgatagattg	1260
tcgcacctga	ttgcccgaca	ttatcgcgag	cccattttata	cccataataa	tcagcatcca	1320
tgttggaatt	taatcgcggc	ctagagcaag	acgtttcccg	ttgaatatgg	ctcataacac	1380
cccttgatatt	actgtttatg	taagcagaca	gttttattgt	tcatgaccac	aatcccttaa	1440
cgtgagtttt	cgttccaactg	agcgtcagac	cccgtagaaa	agatcaaagg	atcttcttga	1500
gatccctttt	ttctgcgcgt	aatctgctgc	ttgcaaacaa	aaaaaccacc	gctaccagcg	1560
gtggtttggt	tgccggatca	agagctacca	actctttttc	cgaaggtaac	tggtctcagc	1620
agagcgcga	taccaaatac	tgtccttcta	gtgtagccgt	agttaggcca	ccacttcaag	1680
aactctgtag	cacgcgctac	atacctcgct	ctgctaattc	tgttaccagt	ggctgctgcc	1740
agtggcgata	agtcgtgtct	tacggggttg	gactcaagac	gatagttacc	ggataaggcg	1800
cagcggtcgg	gctgaaccggg	gggttcgtgc	acacagccca	gcttggagcg	aacgacctac	1860
accgaactga	gatacctaca	gcgtgagcta	tgagaaagcg	ccacgcttcc	cgaagggaga	1920
aaggcggaca	ggtatccggt	aagcggcagg	gtcggaacag	gagagcgcac	gagggagctt	1980
ccagggggaa	acgcctggta	tctttatagt	cctgtcgggt	ttcgccacct	ctgacttgag	2040
cgtcgatttt	tgtgatgctc	gtcagggggg	cggagcctat	ggaaaaacgc	cagcaacgcg	2100
gcctttttac	ggttcctggc	cttttgctgg	ccttttgctc	acatgttctt	tcctgcgtta	2160
tcccctgatt	ctgtggataa	ccgtattacc	gcctttgagt	gagctgatac	cgctcgccgc	2220
agccgaacga	ccgagcgcag	cgagtcagtg	agcgaggaag	cggaagagcg	cctgatgcgg	2280
tattttctcc	ttacgcctct	gtgcggtatt	tcacaccgca	tatatggtgc	actctcagta	2340
caatctgctc	tgatgccgca	tagttaagcc	agtatacact	ccgctatcgc	taogtgactg	2400
ggtcatggct	gcgccccgac	acccgccaac	acccgctgac	gcgcccgtac	gggcttgtct	2460
gctcccggca	tccgcttaca	gacaagctgt	gaccgtctcc	gggagctgca	tgtgtcagag	2520
gttttcaccg	tcatcaccca	aacgcgcgag	gcagctgcgg	taaaagctcat	cagcgtggtc	2580
tggaagcgat	tcacagatgt	ctgcctgttc	atccgcgtcc	agctcgttga	gtttctccag	2640
aagcgttaat	gtctggcttc	tgataaagcg	ggccatgtta	agggcggttt	tttctgtttt	2700
ggtcactgat	gcctcgcgtg	aagggggatt	tctgttcatt	ggggtaatga	taccgatgaa	2760
acgagagagg	atgctcacga	tacgggttac	tgatgatgaa	catgcccggg	tactggaacg	2820
ttgtgagggt	aaacaactgg	cggtatggat	gcggcgggac	cacagaaaaa	tactcaggg	2880
tcaatgccag	cgttcgttta	atacagatgt	aggtgttcca	cagggtagcc	agcagcatcc	2940
tgcatgagc	atccggaaca	taatggtgca	gggcgctgac	ttccgcgttt	ccagacttta	3000
cgaaacacgg	aaaccgaaga	ccattcatgt	tggtgctcag	gtcgcagacg	ttttgcagca	3060
gcagtgcgtt	caggttcgct	cgcgtatcgg	tgattcattc	tgctaaccag	taaggcaacc	3120
ccgccagcct	agccgggtcc	tcaacgacag	gagcacgata	atgcgcaccc	gtggggccgc	3180
catgccggcg	ataatggcct	gcttctcgcc	gaaacgtttg	gtggcgggac	cagtgacgaa	3240
ggcttgagcg	agggcggtga	agattccgaa	taccgcaagc	gacaggccga	tcatcgctgc	3300
gctccagcga	aagcggctcc	cgccgaaaat	gaccagagc	gctgcgggca	cctgtcctac	3360
gagttgcatg	ataaagaaga	cagtcataag	tgcggcgacg	atagtcattg	cccgcgcccc	3420
ccggaaggag	ctgactgggt	tgaaggctct	caagggcatc	ggtcgagatc	ccggtgccta	3480
atgagtgagc	taacttacat	taattgcgtt	gcgtcactg	cccgttttcc	agtcgggaaa	3540
cctgtcgtgc	cagctgcatt	aatgaatcgg	ccaacgcgcg	gggagaggcg	gtttgcgtat	3600
ttgggcccag	ggtgtgtttt	cttttcacca	gtgagacggg	caacagctga	ttgcccttca	3660
ccgcctggcc	ctgagagagt	tgcagcaagc	ggtccacgct	ggtttgcccc	agcaggcgaa	3720
aatcctgttt	gatggtgggt	aacggcggga	tataacatga	gctgtcttcg	gtatcgtcgt	3780
atcccactac	cgagatatcc	gcaccaacgc	gcagcccgga	ctcggtaatg	gcgcgcattg	3840
cgcccagcgc	catctgatcg	ttggcaacca	gcctcgagct	gggaacgatg	ccctcattca	3900
gcatttgcat	ggtttgttga	aaaccggaca	tggcactcca	gtcgcccttc	cgttccqcta	3960

tgcgtgaat	ttgattgcga	gtgagatatt	tatgccagcc	agccagacgc	gacagccgcg	4020
agacagaact	taatgggccc	gctaacagcg	cgatttgctg	gtgacccaat	gcgaccagat	4080
gctccacgcc	cagtcgcgta	cgcgtttcat	gggagaaaat	aatactgttg	atgggtgtct	4140
ggtcagagac	atcaagaaat	aacgcgggaa	cattagtgca	ggcagcttcc	acagcaatgg	4200
catcctggtc	atccagcgga	tagttaatga	tcagcccact	gacgcgttgc	gcgagaagat	4260
tgtgcaccgc	cgctttacag	gcttcgacgc	cgcttcgttc	taccatcgac	accaccacgc	4320
tggcaccag	ttgatcggcg	cgagatttaa	tcgcgcgcac	aattttgcgac	ggcgcgtgca	4380
gggcagact	ggaggtggca	acgccaatca	gcaacgactg	tttgcccgcc	agttgtttgtg	4440
ccacgcggtt	gggaatgtaa	ttcagctccg	ccatcgccgc	ttccactttt	tccgcggttt	4500
tcgcagaaac	gtggctggcc	tggttcacca	cgcgggaaac	ggtctgataa	gagacacggg	4560
catactctgc	gacatcgtat	aacgttactg	gtttcacatt	caccaccctg	aattgactct	4620
cttcggggcg	ctatcatgcc	ataccgcgaa	aggttttgcg	ccattcgatg	gtgtccggga	4680
tctcgacgct	ctcccttatg	cgactcctgc	attaggaagc	agcccagtag	taggttgagg	4740
cggttgagca	ccgccgcgcg	aaggaatggt	gcatgcaagg	agatggcgcc	caacagtocc	4800
ccggccacgg	ggcctgccac	cataccacag	ccgaaacaag	cgctcatgag	cccgaagtgg	4860
cgagcccgat	cttccccatc	ggtgatgtgc	gcgatatagg	cgccagcaac	cgcacctgtg	4920
gcgcgggtga	tgcgcgccac	gatgcgtccg	gcgtagagga	tcgagatctc	gatccgcgca	4980
aattaatacy	actcactata	ggggaattgt	gagcggataa	caattccoct	ctagaataaa	5040
ttttgtttaa	ctttaagaag	gagatataca	tatgcagcat	caccaccatc	accaggagt	5100
acagcttcaa	gacaatgggt	ataatggatt	gctcattgca	attaatcctc	aggtacctga	5160
gaatcagaac	ctcatctcaa	acattaagga	aatgataact	gaagcttcat	tttacctatt	5220
taatgctacc	aagagaagag	tatttttcag	aaatataaag	attttaatac	ctgccacatg	5280
gaaagctaata	aataacagca	aaataaaaca	agaatcatat	gaaaaggcaa	atgtcatagt	5340
gactgactgg	tatggggcac	atggagatga	tccatacacc	ctacaataca	gaggggtgtg	5400
aaaagaggga	aaatacattc	atttcacacc	taatttctta	ctgaatgata	acttaacagc	5460
tggctacgga	tcacgaggcc	gagtgtttgt	ccatgaatgg	gcccacctcc	gttgggggtg	5520
gttcgatgag	tataacaatg	acaaaccttt	ctacataaat	gggcaaaatc	aaattaaagt	5580
gacaaggtgt	tcattctgaca	tcacaggcat	ttttgtgtgt	gaaaaagggtc	cttgccccc	5640
agaaaactgt	attattagta	agcttttttaa	agaaggatgc	accttttatct	acaatagcac	5700
ccaaaatgca	actgcatcaa	taatgttcat	gcaaagttta	tcttctgtgg	ttgaattttg	5760
taatgcaagt	accacaaccc	aagaagcacc	aaacctacag	aaccagatgt	gcagcctcag	5820
aagtgcagtg	gatgtaatca	cagactctgc	tgactttcac	cacagctttc	ccatgaacgg	5880
gactgagctt	ccacctcttc	ccacattctc	gcttgtagag	gctggtgaca	aagtgggtctg	5940
tttagtgcgtg	gatgtgtcca	gcaagatggc	agaggctgac	agactccttc	aactacaaca	6000
agccgcagaa	ttttattttga	tgcagattgt	tgaaattcat	accttcgtgg	gcattggccag	6060
tttcgacagc	aaaggagaga	tcagagccca	gctacaccaa	attaacagca	atgatgatcg	6120
aaagttgctg	gtttcatatc	tgccaccac	tgtatcagct	aaaacagaca	tcagcatttg	6180
ttcagggtct	aagaaaggat	ttgaggtggt	tgaaaaactg	aatggaaaag	cttatggctc	6240
tgtgatgata	ttagtgaacca	gcggagatga	taagcttctt	ggcaattgct	taccactgt	6300
gctcagcagt	ggttcaacaa	ttcactccat	tgccttgggt	tcattctgcag	ccccaaatct	6360
ggaggaatta	tcacgtctta	caggaggttt	aaagtctttt	gttcagata	tatcaaactc	6420
caatagcatg	attgatgctt	tcagtagaat	ttcctctgga	actggagaca	ttttccagca	6480
acataattcag	cttgaaagta	caggtgaaaa	tgtcaaacct	caccatcaat	tgaaaaacac	6540
agtgactgtg	gataatactg	tgggcaacga	cactatgttt	ctagttacgt	ggcaggccag	6600
tggctcctct	gagattatat	tatttgatcc	tgatggacga	aaatactaca	caaataattt	6660
tatcaccaat	ctaacttttc	ggacagctag	tctttggatt	ccaggaacag	ctaagcctgg	6720
gcaactggact	tacaccctga	acaataccca	tcattctctg	caagccctga	aagtgcagat	6780
gacctctcgc	gcttccaact	cagctgtgcc	cccagccact	gtggaagcct	ttgtggaaag	6840
agacagcttc	cattttcttc	atcctgtgat	gatttatgcc	aatgtgaaac	agggatttta	6900
tcccatcttt	aatgccactg	tcaactgccac	agttgagcca	gagactggag	atcctgttac	6960
gctgagactc	cttgatgatg	gagcaggtgc	tgatgttata	aaaaatgatg	gaatttactc	7020
gaggtatttt	ttctcctttg	ctgcaaatgg	tagatatagc	ttgaaagtgc	atgtcaatca	7080
ctctcccagc	ataagcacc	cagccactc	tattccaggg	agtcatgcta	tgtatgtacc	7140
aggttacaca	gcaaacggta	atattcagat	gaatgctcca	aggaatcag	taqqcagaaa	7200

```

tgaggaggag cgaaaagtggg gcttttagccg agtcagctca ggaggctcct tttcagtgc 7260
gggagttcca gctggccccc accctgatgt gtttccacca tgcaaaatta ttgacctgga 7320
agctgtaaaa gtagaagagg aattgacct atcttggaac gcacctggag aagactttga 7380
tcaggggccag gctacaagct atgaaataag aatgagtaaa agtctacaga atatccaaga 7440
tgactttaac aatgctattt tagtaaatac atcaaagcga aatcctcagc aagctggcat 7500
caggagata tttacgttct caccocaaat ttccacgaat ggacctgaac atcagccaaa 7560
tgagaaaaca catgaaagcc acagaattta tgttgcaata cgagcaatgg ataggaactc 7620
cttacagtct gctgtatcta acattgcccc ggccgctctg tttattcccc ccaattctga 7680
tcctgtacct gccagagatt atcttatatt gaaaggagtt ttaacagcaa tgggtttgat 7740
aggaatcatt tgccttatta tagttgtgac acatcatact ttaagcagga aaaagagagc 7800
agacaagaaa gagaatggaa caaaattatt ataatgaatt ctgcagatat ccatcacact 7860
ggcggccgct cgagcaccac caccaccacc actgagatcc ggctgctaac aaagcccgaa 7920
aggaagctga gttggtctgt gccaccgctg agcaataact agcataaccc cttggggcct 7980
ctaaacgggt cttgaggggt tttttgctga aaggaggaac tatatccgga t 8031

```

```

<210> 255
<211> 401
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(401)
<223> n = A,T,C or G

```

```
<400> 255
```

```

gtggccagng actagaaggg gagggcgcgc gggaccatgg cggcggcggc ggacgagcgg 60
agtccanagg acggagaaga cgaggaagag gaggagcagt tggttctggt ggaattatca 120
ggaattattg attcagaactt cctctcaaaa tgtgaaaata aatgcaagggt tttgggcatt 180
gacactgaga ggcccattct gcaagtggac agctgtgtct ttgctgggga gtatgaagac 240
actctangga cctgtgttat atttgaagaa aatgntnaac atgctgatac agaaggcaat 300
aataaaacag tgctaaaata taaatgccat acaatgaaga agctcagcat gacaagaact 360
ctcctgacag agaagaagga aggagaagaa aacatangtg g 401

```

```

<210> 256
<211> 401
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(401)
<223> n = A,T,C or G

```

```
<400> 256
```

```

tggtggncc tggatgggga accgcggtgg cttccgnnga ggtttcggca ntggcatccg 60
gggcccgggt cgcgcccgng gacggggcgc gggccnangc cgnnganctc gcggangcaa 120
ggccgaggat aaggagtggg tgcccgtcac caacttgggc cgttgncca aggacatgaa 180
nancaaagccc ctgnaggaga tctatntctt cttccctgcc ccattaagga atcaagagat 240
catttgattt ctccctgggg gcctctctca aggatnaggt ttttgaagat tatgccagtg 300
canaaannan accccgttgc cngtccatc tncacccaac ncttccaagg gcnatttttg 360
tttaggcctc attncngggg ggaaccttaa cccaatttgg g 401

```

<210> 257
 <211> 401
 <212> DNA
 <213> Homo sapien

 <220>
 <221> misc_feature
 <222> (1)...(401)
 <223> n = A,T,C or G

<400> 257
 atgtatgtaa aacacttcat aaaatgtaaa gggctataac aaatatgtta taaagtgatt 60
 ctctcagccc tgaggtatac agaatcattt gcctcagact gctgttgat tttaaaattt 120
 ttaaaatata tgctaagtaa tttgctatgt cttctccac actatcaata tgcctgcttc 180
 taacaggctc cccactttct tttaatgtgc tgttatgagc tttggacatg agataaccgt 240
 gcctgttcag agtgtctaca gtaagagctg gacaaactct ggaggacac agtctttgag 300
 acagctcttt tgggtgcttt ccacttttct gaaagggttca cagtaacctt ctagataata 360
 gaaactccca gttaaagcct angtancaa ttttttttag t 401

<210> 258
 <211> 401
 <212> DNA
 <213> Homo sapien

<400> 258
 ggagcgctag gtcggtgtac gaccgagatt aggggtgcgtg ccagctccgg gaggccgcgg 60
 tgaggggccc ggcccaagct gccgacccga gccgatcgtc agggctcgcca gcgcctcagc 120
 tctgtggagg agcagcagta gtcggagggt gcaggatatt agaaatggct actccccagt 180
 caattttcat ctttgcaatc tgcattttta tgataacaga attaattctg gcctcaaaaa 240
 gctactatga tatcttaggt gtgccaaaat cggcatcaga gcgccaaatc aagaaggcct 300
 ttcacaagtt ggccatgaag taccacccctg acaaaaaata gaccagatg ctgaagcaaa 360
 attcagagag attgcagaag catatgaaac actctcagat g 401

<210> 259
 <211> 401
 <212> DNA
 <213> Homo sapien

<400> 259
 attgggtttg gaggaggat gatgacagag gaatgccctt tggccatcac ggttttgatt 60
 ctccagaata ttgtgggttt gatcatcaat gcagtcatgt taggctgcat tttcatgaaa 120
 acagctcagg ctacagaag ggcagaaact ttgattttca gccgccatgc tgtgattgcc 180
 gtccgaaatg gcaagctgtg cttcatgttc cgagtgggtg acctgaggaa aagcatgac 240
 attagtgcct ctgtgcgcat ccagggtggtc aagaaaacaa ctacacctga aggggaggtg 300
 gttcctatcc accaactgga cattcctgtt gataacccaa tcgagagcaa taacattttt 360
 ctggtggccc ctttgatcat ctgccacgtg attgacaagc g 401

<210> 260
 <211> 363
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature

<222> (1)...(363)

<223> n = A,T,C or G

<400> 260

```
aggaganang gaggggggana tgaatagggg tggagagggg natagtggat gagcagggca      60
canggagagg aancagaaag gagaggcaag acagggagac acacancaca nangangana      120
caggtggggg ctggggtggg gcatggagag ccttttnangt cccccaggcc accctgctct      180
cgctggnctg ttgaaaccca ctccatggct tcctgccact gcagttgggc ccagggctgg      240
cttattnctg gaatgcaagt ggctgtggct tggagcctcc cctctggnnn anggaaannn      300
attgctccct tatctgcttg gaatatctga gtttttccan cccggaaata aaacacacac      360
aca                                                                    363
```

<210> 261

<211> 401

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(401)

<223> n = A,T,C or G

<400> 261

```
cggctctccg ccgctctccc ggggtttcgg ggcaacttggg tcccacagtc tggtcctgct      60
tcaccttccc ctgacctgag tagtcgccat ggcacagggt ctcagaggca ctgngactga      120
cttccttgga tttgatgagc gggctgatgc anaaactctt cggaaggcta tgaaaggctt      180
gggcacagat gaggagagca tcctgactct gttgacatcc cgaagtaatg ctcagcgcca      240
ggaaatctct gcagctttta agactctggt tggcagggat cttctggatg acctgaaatc      300
agaactaact ggaaaatttg aaaaattaat tgtggctctg atgaaaccct ctcggtttta      360
tgatgcttat gaactgaaac atgccttgaa gggagctgga a                                                                    401
```

<210> 262

<211> 401

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(401)

<223> n = A,T,C or G

<400> 262

```
agtctanaac atttctaata ttttngctt tcatatatca aaggagatta tgtgaaacta      60
tttttaataa ctgtaaagtg acatatagtt ataagatata tttctgtaca gtagagaaag      120
agtttataac atgaagaata ttgtaccatt atacattttc attctcgatc tcataagaaa      180
ttcaaaagaa taatgataga ggtgaaaata tgtttacttt ctctaaatca agcctagttg      240
tcaactcaaa aattatgntg catagtttta ttttgaattt aggttttggg actacttttt      300
tccancttca atgagaaaaa aaaatctaca actcaggagt tactacagaa gttotaanta      360
tttttttgct aannagcnaa aaatataaac atatgaaaaa g                                                                    401
```

<210> 263

<211> 401

<212> DNA

<213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(401)
 <223> n = A,T,C or G

<400> 263
 ctgtccgacc aagagaggcc ggccgagccc gaggcttggg cttttgcttt ctggcggagg 60
 gatctgcggc ggtttaggag gcggcgctga tcctgggagg aagaggcagc tacggcggcg 120
 gcggcgggtg cggctagggc ggccgcgaat aaaggggccc ccgccgggtg atgcggtgac 180
 cactgcggca ggcccaggag ctgagtgggc cccggccctc agcccgtccc gncggaccgc 240
 ctttcctcaa ctctccatct tctcctgccg accgagatcg ccgaggcggn ctcaggctcc 300
 ctanccctt ccccgctccct tccccncccc cgcccccgcc ccggggggccg ccgccaccgc 360
 cctcccacca tggctctgaa ganaatccac aaggaattga a 401

<210> 264
 <211> 401
 <212> DNA
 <213> Homo sapien

<400> 264
 aacaccagcc actccaggac ccctgaaggc ctctaccagg tcaccagtgt tctgcgccta 60
 aagccacccc ctggcagaaa cttcagctgt gtgtttctgga ataactcacgt gagggaaactt 120
 actttggcca gcattgacct tcaaagtcag atggaacca ggaccatcc aacttggtg 180
 cttcacattt tcatccctc ctgcatcatt gctttcattt tcatagccac agtgatagcc 240
 ctaagaaaac aactctgtca aaagctgtat tcttcaaaa acacaacaaa aagacctgtc 300
 accacaacaa agaggggaagt gaacagtgt gtgaatctga acctgtggtc ttgggagcca 360
 ggggtgacctg atatgacatc taaagaagct tctggactct g 401

<210> 265
 <211> 271
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(271)
 <223> n = A,T,C or G

<400> 265
 gccacttcct gtggacatgg gcagagcgct gctgccagtt cctggtagcc ttgaccacna 60
 cgctgggggg tctttgtgat ggtcatgggt ctcatttgca cttgggggtg tgggattcaa 120
 gttagaagtt tctagatctg gccgggcgca gtggctcaca cctgtaatcc cagcacttta 180
 ggaggctgag gcaggcggt catgaggtca ggagatcgag accgtcctgg ctaacacagt 240
 gaaaccccgct ctctactaaa aatacaaaaa a 271

<210> 266
 <211> 401
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(401)

<223> n = A,T,C or G

<400> 266

```
attcataaat ttagctgaaa gatactgatt caatttgtat acagngaata taaatgagac      60
gacagcaaaa ttttcatgaa atgtaaaata tttttatagt ttgttcatac tatatgaggt    120
tctattttaa atgactttct ggatttttaa aaatttcttt aaatacaatc atttttgtaa    180
tatttatatt atgottatga tctagataat tgcagaatat cattttatct gactctgtct    240
tcataagaga gctgtggccg aattttgaac atctgttata gggagtgatc aaattagaag    300
gcaatgtgga aaaacaattc tgggaaagat ttctttatat gaagtccttg ccactagcca    360
gccatcctaa ttgatgaaag ttatctgttc acaggcctgc a                          401
```

<210> 267

<211> 401

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(401)

<223> n = A,T,C or G

<400> 267

```
gaagaggcat cacctgatcc cggagacett tggagttaag aggcggcgga agcgagggcc      60
tgtggagtcg gatcctcttc ggggtgagcc agggctcgcg cgcgcggtg tctcanaact    120
catgcagctg ttcccgcgag gcctgtttga ggacgcgctg ccgcccatcg tgctgaggag    180
ccaggtgtac agccttgtgc ctgacaggac cgtggccgac cggcagctga aggagcttca    240
agagcanggg gagacaaaat cgtccagctg ggcttcnact tggatgcca tggaanttat    300
tctttcnctt ganggactta cnnnggaccc aagaancctt tncaaggggc ccttngtgga    360
tgggncccga aaccccnnta tttgcccttg ggggggncca a                          401
```

<210> 268

<211> 223

<212> DNA

<213> Homo sapien

<400> 268

```
tgcctatgtt ggccaggctg gtcttgaact cctgacttta agtgatccac ccgcctcaac      60
ctcccaaagt gctgggatta cagggtgtgag ccaccgcgcc tggcctgata catactttta    120
gaatcaagta gtcacgcaact ttttctgttc atttttctaa aaagtaaata tacaaatgtt    180
ttgttttttg ttttttttgt ttgtttgttt ctgttttttt ttt                      223
```

<210> 269

<211> 401

<212> DNA

<213> Homo sapien

<400> 269

```
actatgtaaa ccacattgta ctttttttta ctttggcaac aaatatttat acatacaaga      60
tgctagtcca tttgaatatt tctcccaact tatccaagga tctccagctc taacaaaatg    120
gtttattttt atttaaattg caatagtgtg tttttaaaat ccaaatcaga ggtgcaggcc    180
accagttaaa tgccgtctat cagggttttg gccttaagag actacagagt caaagctcat    240
ttttaaagga gtaggacaaa gttgtcacag gttttgttg ttgtttttat tgcccccaaa    300
attacatgtt aatttcatt tatatcaggg attctattha cttgaagact gtgaagttgc    360
cattttgtct cattgttttc tttgacataa ctaggatcca t                          401
```

```
<220>
<221> misc_feature
<222> (1)...(401)
<223> n = A,T,C or G
```

```
<210> 271
<211> 329
<212> DNA
<213> Homo sapien
```

```
<210> 272
<211> 401
<212> DNA
<213> Homo sapien
```

```
<220>
<221> misc_feature
<222> (1)...(401)
<223> n = A,T,C or G
```

<210> 273
<211> 401

```
<220>
<221> misc_feature
<222> (1)...(401)
<223> n = A,T,C or G
```

```
<210> 274
<211> 401
<212> DNA
<213> Homo sapien
```

```
<210> 275
<211> 401
<212> DNA
<213> Homo sapien
```

```
<210> 276
<211> 401
<212> DNA
<213> Homo sapien
```

```
<220>
<221> misc feature
<222> (1)...(401)
<223> n = A,T,C or G
```

<400> 276

```
tctgatattg ntacccttga gccacctaag ttagaagaaa ttggaaatca agaagttgtc      60
attggttgaag aagcacagag ttcagaagac ttttaacatgg gctcttcctc tagcagccag    120
tatactttct gtcagccaga aactgtatct tcatctcagc ctagtgatga tgaatcaagt    180
agtgatgaaa ccagtaatca gcccagtcct gccttttagac gacgccgtgc taggaagaag    240
accgtttctg cttcagaatc tgaagaccgg ctagttgggtg aacaagaaac tgaaccttct    300
aaggagttga gtaaactgca gttcagtagt ggtctcaata agtgtgttat acttgctttg    360
gtgattgcaa tcagcatggg atttggccat ttctatggca c                          401
```

<210> 277

<211> 401

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(401)

<223> n = A,T,C or G

<400> 277

```
aactttggca acatatctca gcaaaaacta cagctatggt attcatgcc aataaaaagc      60
tgtgcagagg agtggctgca atgaggtcac aacgggtggg gatgtaaaag agatcttcaa    120
gtcctcatca cccatccctc gaactcaagt cccgctcatt acaaatctct cttgccagtg    180
tccacacatc ctgccccatc aagatgttct catcatgtgt tacgagnggc gctcaaggat    240
gatgcttctt gaaaattgct tagttgaaaa atggagagat cagcttagta aaagatccat    300
acagtgggaa gagaggctgc aggaacagcg ganaacagtt caggacaaga agaaaacagc    360
cgggcgcacc agtcgtagta atccccccaa accaaaggga a                          401
```

<210> 278

<211> 401

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(401)

<223> n = A,T,C or G

<400> 278

```
aatgagtgtg agaccacaaa tgaatgccgg gaggatgaaa tgtgttgga ttatcatggc      60
ggcttccgtt gttatccacg aaatccttgt caagatccct acattctaac accagagaac    120
cgatgtgttt gccagtcctc aaatgccatg tgccgagaaac tgccccagtc aatagtctac    180
aaatacatga gcatccgatc tgataggctc gtgccatcag acatcttcca gatacaggcc    240
acaactatct atgccaacac catcaatact ttccggatta aatctggaaa tgaaaatgga    300
gagtctacct acgacaacaa anccctgtaa gtgcaatgct tgtgctcgtg aagncattat    360
caggaccaag agaacatatc gtggacctgg agatgctgac a                          401
```

<210> 279

<211> 401

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature
 <222> (1)...(401)
 <223> n = A,T,C or G

<400> 279

aaattattgc	ctctgataca	tacctaagtn	aacanaacat	taatacctaa	gtaaacataa	60
cattacttgg	agggttgcag	nttctaantg	aaactgtatt	tgaaactttt	aagtatactt	120
taggaaacaa	gcatgaacgg	cagtctagaa	taccagaaac	atctacttgg	gtagcttggn	180
gccattatcc	tgtggaatct	gatatgtctg	gnagcatgtc	attgatggga	catgaagaca	240
tctttgaaa	tgatgagatt	atttcctgtg	ttaaaaaaaa	aaaaaatctt	aaattcctac	300
aatgtgaaac	tgaaactaat	aattttgatc	ctgatgtatg	ggacagcgta	tctgtaccag	360
gctctaaata	acaaaagnta	gggngacaag	nacatgttcc	t		401

<210> 280
 <211> 326
 <212> DNA
 <213> Homo sapien

<400> 280

gaagtggaat	tgtataattc	aattcgataa	ttgatctcat	gggctttccc	tggaggaaag	60
gttttttttg	ttgttttttt	tttaagaact	tgaaacttgt	aaactgagat	gtctgtagct	120
tttttgccca	tctgtagtgt	atgtgaagat	ttcaaaacct	gagagcactt	tttctttggt	180
tagaattatg	agaaaggcac	tagatgactt	taggatttgc	atttttccct	ttattgcctc	240
atttcttggtg	acgccttggt	ggggagggaa	atctgtttat	tttttccctac	aaataaaaaag	300
ctaagattct	atatcgcaaa	aaaaaa				326

<210> 281
 <211> 374
 <212> DNA
 <213> Homo sapien

<400> 281

caacgcgttt	gcaaattatc	ccttggtagc	ctacttccct	acccccgaat	attggtaaga	60
tcgagcaatg	gcttcaggac	atgggttctc	ttctcctgtg	atcattcaag	tgctcactgc	120
atgaagactg	gcttgtctca	gtgtttcaac	ctcaccagg	ctgtctcttg	gtccacacct	180
cgctccctgt	tagtgccgta	tgacagcccc	catcaaatga	ccttgggcaa	gtcacgggtt	240
ctctgtggtc	aagggtgggt	ggctgattgg	tggaaagttag	ggtggacca	aggaggccac	300
gtgagcagtc	agcaccagtt	ctgcaccagc	agcgcctccg	tcctagtggg	tgttcctggt	360
tctcctggcc	ctgg					374

<210> 282
 <211> 404
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(404)
 <223> n = A,T,C or G

<400> 282

agtgtggtgg	aattcccgc	tctanncg	cgactcacac	aaggcagagt	ngccatggag	60
aaaattccag	tgtcagcatt	cttgctcctt	gtggccctct	cctacactct	ggccagagat	120
accacagtca	aacctgnagc	caaaaaggac	acaaaaggact	ctcgacccaa	actgccccan	180

```

accctctcca gaggttgggg tgaccaactc atctggactc anacatatga agaagctcta 240
tataaatcca agacaagcaa caaaccttg atgattattc atcacttgga tgagtgccca 300
cacagtcaag ctttaaagaa agtgtttgct gaaaataaag aaatccagaa attggcagag 360
cagtttgctc tcctcaatct ggtttatgaa acaactgaca aaca 404

```

```

<210> 283
<211> 184
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(184)
<223> n = A,T,C or G

```

```

<400> 283
agtgtggtgg aattcacttg cttaanttgt gggcaaaaga gaaaaagaag gattgatcag 60
agcattgtgc aatacagttt cattaactcc ttccctcgct cccccaaaaa tttgaatttt 120
ttttcaaca ctcttacacc tggtatggaa aatgtcaacc tttgtaagaa aaccaaaata 180
aaaa 184

```

```

<210> 284
<211> 421
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(421)
<223> n = A,T,C or G

```

```

<400> 284
ctattaatcc tgccacaata tttttaatta cgtacaaaga tctgacatgt caccagggga 60
ccatttcac ccactgctct gtttggccgc cagtcttttg tctctctctt cagcaatggg 120
gaggcgata ccttttcctc ggggaanana aatccatggt ttgttgccct tgccaataac 180
aaaaatgttg gaaagtcgag tggcaaagct gttgccattg gcatctttca cgtgaaccac 240
gtcaaaagat ccagggtgcc tctctctggt ggtgatcaca ccaattcttc ctaggtttagc 300
acctccagtc accatacaca gggtaccagt gtcgaacttg atgaaatcag taatcttgcc 360
agtctctaaa tcaatctgaa tggatcatt caccttgatg aggggatcgg ggtagcggat 420
g 421

```

```

<210> 285
<211> 361
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(361)
<223> n = A,T,C or G

```

```

<400> 285
ctgggtggta actctttatt tcattgtccg gaanaaagat gggagtggga acagggtgga 60
cactgtgcag gcttcagctt ccactccggg caggattcag gctatctggg accgcaggga 120

```

```

ctgccagggtg cacagccctg gctcccagag caggcaggca aggtgacggg actggaagcc 180
cttttcanag ccttgaggga gctgggtccg ccacaagcaa tgagtgccac tctgcagttt 240
gcaggggatg gataaacagg gaaacactgt gcattcctca cagccaacag tgtaggtctt 300
ggtgaagccc cggcgctgag ctaagctcag gctgttccag ggagccacga aactgcaggt 360
a 361

```

```

<210> 286
<211> 336
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(336)
<223> n = A,T,C or G

```

```

<400> 286
tttgagtggc agcgcttta tttgtggggg ccttcaaggn agggtcgtgg ggggcagcgg 60
ggaggaanag ccganaaact gtgtgaccgg ggccctcagg ggtgggcatt gggggctcct 120
cttgcanatg cccattggca tcaccgggtg agccattggg ggcagcgggt accggtcctt 180
tcttgttcaa catagggtag gtggcagcca cgggtccaac tcgcttgagg ctgggccctg 240
ggcgctccat tttgtgttcc angagcatgt ggttctgtgg cgggagcccc acgcaggccc 300
tgaggatgtt ctcgatgcag ctgcgctggc ggaaaa 336

```

```

<210> 287
<211> 301
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(301)
<223> n = A,T,C or G

```

```

<400> 287
tgggtaccaaa atttntttat ttgaaggaat ggnacaaatc aaanaactta agnggatgtt 60
ttggtacaac ttatanaaaa ggnaaaggaa accccaacat gcatgcnctg ccttgngnac 120
caggaagtc accccaagggc tatggggaaa ttancccgag gcttancctt cattatcact 180
gtctcccagg gngngcttgt caaaaanata ttcnccaag ccaaattcgg gcgctcccat 240
nttgcnaag ttggtcacgt ggtcacccaa ttctttgatg gctttcacct gctcattcag 300
g 301

```

```

<210> 288
<211> 358
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(358)
<223> n = A,T,C or G

```

```

<400> 288
aagtttttaa actttttatt tgcatattaa aaaaattgng cattccaata attaaatca 60

```



```
<210> 289
<211> 462
<212> DNA
<213> Homo sapien
```

<400> 289

```
<210> 290
<211> 481
<212> DNA
<213> Homo sapien
```

<400> 290

```
<210> 291
<211> 381
<212> DNA
<213> Homo sapien
```

```
<220>  
<221> misc feature
```

<222> (1)...(381)

<223> n = A,T,C or G

<400> 291

```
tcataagtaat gtaaaacccat ttgtttaatt ctaaatacaaa tcactttcac aacagtgaag 60
attagtgaact ggtaagng tgccactgta catatcatca tttctgact ggggtcagga 120
cctggctcta gtccacaagg gtggcaggag gaggggtggag gctaanaaca cagaaaacac 180
acaaaanaaa ggaaagctgc cttggcanaa ggatgagng gtgagcttgc cgaaggatgg 240
tggaagggg gctccctgtt ggggcccagc caggagtccc aagtcagctc tctgcctta 300
cttagctcct ggcanagggt gagtggggac ctacgaggtt caaaatacaaa tggcatttgg 360
ccagcctggc ttactaaca g 381
```

<210> 292

<211> 371

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(371)

<223> n = A,T,C or G

<400> 292

```
gaaaaaataa tccgtttaat tgaaaaacct gnaggataact attccactcc cccanattgag 60
gaggctgagg anaccaaacc cctacatcac ctctgtagcca cttctgatac tcttcacgag 120
gcagcaggca aagacaattc ccaaaacctc nacaaaagca attccaaggg ctgctgcagc 180
taccaccanc acatttttcc tcagccagcc cccaatcttc tccacacagc cctccttatg 240
gatcgcttcc tctgtgaaat taatcccaca gccacagta acattaatgc ancaggagtc 300
ggggactcgg ttcttcgaca tggaagggat tttctcccaa tctgtgtagt tagcagcccc 360
acagcactta a 371
```

<210> 293

<211> 361

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(361)

<223> n = A,T,C or G

<400> 293

```
gatttaaaag aaaacacttt attgttcagc aattaaaagt tagccaaata tgtatttttc 60
tccataattt attngatgt tatcaacatc aagtaaaatg ctcattttca tcatttgctt 120
ctgttcatgt tttcttgaac acgtcttcaa tttccttcc aaaatgctgc atgccacact 180
tgaggtaacg aagcanaagt atttttaaac atgacagcta anaacattca tctacagcaa 240
cctatatgct caatacatgc cgcgtgatcc tagtagtttt ttcacaacct tctacaagtt 300
tttgaaaaac atctgttatg atgactttca tacaccttca cctcaaaggc tttcttgcac 360
c 361
```

<210> 294

<211> 391

<212> DNA

<213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(391)
 <223> n = A,T,C or G

<400> 294
 tatttttaaag tttaattatg attcanaaaaa aatcgagcga ataactttct ctgaaaaaat 60
 atattgactc tgtatanacc acagttattg gggganaagg gctggtaggt taaattatcc 120
 tattttttat tctgaaaatg atattaatan aaagtcccg ttcagtcctg attataaaga 180
 tacatatgcc caaaatggct ganaataaat acaacaggaa atgcaaaagc tgtaaagcta 240
 agggcatgca ananaaaatc tcanaataacc caaagnggca acaaggaacg tttggctgga 300
 atttgaagtt atttcagtca tctttgtctt tggtccatg tttcaggatg cgtgtgaact 360
 cgatgtaatt gaaattcccc tttttatcaa t 391

<210> 295
 <211> 343
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(343)
 <223> n = A,T,C or G

<400> 295
 ttcttttggt ttattgataa cagaaactgt gcataattac agatttgatg aggaatctgc 60
 aaataataaa gaatgtgtct actgccagca aaatacaatt attccatgcc ctctcaacat 120
 acaaatatag agttcttcac accanatggc tctgggtgtaa caaagccatt ttanatgttt 180
 aattgtgctt ctacaaaacc ttcanagcat gaggtagttt cttttaccta cnatattttc 240
 cacatttcca ttattacact tttagtgage taaaatcctt ttaacatagc ctgcggatga 300
 tctttcacaa aagccaagcc tcatttacaagggtttatt tct 343

<210> 296
 <211> 241
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(241)
 <223> n = A,T,C or G

<400> 296
 ttcttgata ttggttgttt ttgtgaaaaa gtttttggtt ttcttctcag tcaactgaat 60
 tatttctcta ctttgccctc ctgatgccca catgananaa cttaanataa tttctaacag 120
 cttccacttt ggaaaaaaa aaaacctgtt ttctctatgg aaccccagga gttgaaagtg 180
 gatanatcgc tctcaaaaatc taaggctctg ttcagcttta cattatgtta cctgacgttt 240
 t 241

<210> 297
 <211> 391
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(391)
 <223> n = A,T,C or G

<400> 297

gttggtggctg	anaatgctgg	agatgctcag	ttctctccct	cacaaggtag	gccacaaatt	60
cttggtgggtg	ccctcacatc	tggggctctc	aggcaccagc	catgcctgcc	gaggagtgtc	120
gtcaggacan	accatgtccg	tgctaggccc	aggcacagcc	caaccactcc	tcatccaagt	180
ctctcccagg	tttctgggtc	cgatgggcaa	ggatgacccc	tccagtggct	ggtaccccac	240
catcccacta	ccctcacat	gctctcactc	tccatcaggt	ccccaatcct	ggcttccctc	300
ttcacgaact	ctcaaaagaa	aggaaggata	aaacctaatt	aaaccagaca	gaagcagctc	360
tggaaaagta	caaaaagaca	gccagagggtg	t			391

<210> 298
 <211> 321
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(321)
 <223> n = A,T,C or G

<400> 298

caagccaaac	tgtntccagc	tttatttaaan	atactttcca	taaacaatca	tggtatttca	60
ggcaggacat	gggcanacaa	tcgttaacag	tatacaacaa	ctttcaaact	cccttnttca	120
atggactacc	aaaaatcaaa	aagccactat	aaaacccaat	gaagtcttca	tctgatgtct	180
tgaacaggga	aagttttaaag	ngaggggttg	catttcacat	ttagcatgtt	gtttaacaac	240
ttttcacaag	ccgaccctga	ctttcaggaa	gtgaaatgaa	aatggcanaa	tttatctgaa	300
natccacaat	ctaaaaatgg	a				321

<210> 299
 <211> 401
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(401)
 <223> n = A,T,C or G

<400> 299

tatcataaag	agtgttgaag	tttatttatt	atagcaccat	tgagacattt	tgaaattgga	60
attggtaaaa	aaataaaaca	aaaagcattt	gaattgtatt	tggnggaaca	gcaaaaaaag	120
agaagtatca	tttttctttg	tcaaattata	ctgtttccaa	acatttttga	aataaataac	180
tggaattttg	tcggtcactt	gcactgggtg	acaagattag	aacaagagga	acacatatgg	240
agttaaattt	tttttggttg	gatttcanat	agagtttgg	ttataaaaag	caaacagggc	300
caacgtccac	accaaattct	tgatcaggac	caccaatgtc	ataggnggca	atatctacaa	360
taggtagtct	cacagccttg	cgtgttcgat	attcaaagac	t		401

<210> 300
 <211> 188

<212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(188)
 <223> n = A,T,C or G

<400> 300
 tgaatgcttt gtcataataa gaaagttaaa gtgcaataat gtttgaanac aataagtgg 60
 ggtgtatctt gtttctaata agataaactt ttttgtcttt gctttatctt attagggagt 120
 tgtatgtcag tgtataaaac atactgtgtg gtataacagg cttataaat tctttaaaag 180
 gaaaaaaa 188

<210> 301
 <211> 291
 <212> DNA
 <213> Homo sapien

<400> 301
 aagattttgt tttattttat tatggctaga aagacactgt tatagccaaa atcggcaatg 60
 aactaaaga aatcctctgt gcttttcaat atgcaaata atttcttcca agagttgccc 120
 tgggtgtgact tcaagagttc atgttaactt cttttctgga aacttccttt tcttagttgt 180
 tgtattcttg aagagcctgg gccatgaaga gcttgccata gttttgggca gtgaactcct 240
 tgatgttctg gcagtaagtg tttatctggc ctgcaatgag cagcgagtcc a 291

<210> 302
 <211> 341
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(341)
 <223> n = A,T,C or G

<400> 302
 tgatTTTTCA taattttatt aaatnatcac tgggaaaact aatggttcgc gtatcacaca 60
 attacactac aatctgatag gagtggtaaa accagccaat ggaatccagg taaagtacaa 120
 aaacgccacc ttttattgtc ctgtcttatt tctcgggaag gagggttcta ctttacacat 180
 ttcattgagcc agcagtggac ttgagttaca atgtgtaggt tcttctgtgt tatagctgca 240
 gaagaagcca tcaaattctt gaggacttga catctctcgg aaagaagcaa actagtggat 300
 cccccgggct gcaggaattc gatatcaagc ttatcgatac c 341

<210> 303
 <211> 361
 <212> DNA
 <213> Homo sapien

<220>
 <221> misc_feature
 <222> (1)...(361)
 <223> n = A,T,C or G

<400> 303

tcgacagagt	aatnaat	ttttgngtt	cacagaacat	actaggcgat	ctcgacagtc	60
gctccgtgac	agcccaccaa	cccccaaccc	tntacctcgc	agccacccta	aaggcgactt	120
caanaanatg	gaaggatctc	acggatctca	ttcctaattg	tccgccgaag	tctcacacag	180
tanacagacg	gagttganat	gctggaggat	gcagtcacct	cctaaactta	cgaccaccca	240
ccanacttca	tcccagccgg	gacgtctctc	cccacccgag	tcttccccat	ttcttctctt	300
actttgccgc	agttccaggn	gtctgtcttc	caccagtccc	acaaagctca	ataaatacca	360
a						361

<210> 304

<211> 301

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(301)

<223> n = A,T,C or G

<400> 304

ctctttacaa	cagcctttat	ttncggccct	tgatctctgct	cggatgctgg	tggaggccct	60
tagctccgcc	cgccaggtct	tgtgccgcct	ccccgcaggc	gcanattcat	gaacacgggt	120
ctcaggggct	tgaggccgta	ctccccccagc	gggagctggg	cctccagggg	cttccccctcg	180
aaggtcagcc	anaacaggtc	gtctctgcaca	ccctccagcc	cgctcacttg	ctgcttcagg	240
tggggccacgg	tctgcgtcag	ccgcacctcg	taggtgctgc	tgccggccctt	gttatctctc	300
a						301

<210> 305

<211> 331

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(331)

<223> n = A,T,C or G

<400> 305

ganaggctag	taacatcagt	tttattgggt	tggggnggca	accatagcct	ggctgggggn	60
ggggctggcc	ctcacagggt	gttgagttcc	agcagggtct	ggtccaagg	ctggtgaatc	120
tcgacgttct	cctccttggc	actggccaag	gtctcttcta	ggtcatcgat	ggttttctcc	180
aactttgcc	canacctctc	ggcaaactct	gctcgggtct	cancctcctt	cagcttctcc	240
tccaacaggt	tgatctctct	ttcatattta	tcttcttttg	gggaatactc	ctcctctgag	300
gccatcaggg	acttgagggc	ctggtccatg	g			331

<210> 306

<211> 457

<212> DNA

<213> Homo sapien

<400> 306

aatatgtaaa	ggtaataact	tttattatat	taaagacaat	gcaaacgaaa	aacagaattg	60
agcagtgcaa	aattttaaagg	actgttttgt	tctcaaagtt	gcaagtttca	aagccaaaag	120
aattatatgt	atcaaatata	taagtaaaaa	aaagtttagac	tttcaagcct	gtaatcccag	180

cactttggga	ggctgaggca	ggtggatcac	taacattaaa	aagacaacat	tagattttgt	240
cgatttatag	caattttata	aatatataac	tttgtcactt	ggatcctgaa	gcaaaataat	300
aaagtgaatt	tgggattttt	gtacttggtg	aaaagtttaa	caccctaaat	tcacaactag	360
tggatcccc	gggctgcagg	aattcgatat	caagcttata	gataccgtcg	acctcgaggg	420
ggggcccggt	acccaattcg	ccctatagtg	agtcgta			457

<210> 307

<211> 491

<212> DNA

<213> Homo sapien

<400> 307

gtgcttgga	ggaacccggc	gctcgttccc	caccccggcc	ggccgcccac	agccagccct	60
ccgtcacctc	ttcacgcac	cctcggactg	ccccaggcc	cccgccgccc	ctccagcgcc	120
gcgagccac	cgccgcgccc	gccgcctctc	cttagtcgcc	gccatgacga	ccgcgtccac	180
ctcgaggtg	cgccagaact	accaccagga	ctcagaggcc	gccatcaacc	gccagatcaa	240
cctggagctc	tacgcctcct	acgtttacct	gtccatgtct	tactactttg	accgcgatga	300
tgtggctttg	aagaactttg	caaatacttt	tcttcaccaa	tctcatgagg	agagggaaca	360
tgctgagaaa	ctgatgaagc	tgcagaacca	acgaggtggc	cgaatcttcc	ttcaggatat	420
caagaaacca	gactgtgatg	actgggagag	cgggctgaat	gcaatggagt	gtgcattaca	480
tttgaaaaa	a					491

<210> 308

<211> 421

<212> DNA

<213> Homo sapien

<400> 308

ctcagcgctt	cttctttctt	ggtttgatcc	tgactgctgt	catggcgtgc	cctctggaga	60
aggccctgga	tgtgatggtg	tccaccttcc	acaagtactc	gggcaaagag	ggtgacaagt	120
tcaagctcaa	caagtcagaa	ctaaaggagc	tgtgaccccg	ggagctgccc	agcttcttgg	180
ggaaaaggac	agatgaagct	gctttccaga	agctgatgag	caacttgga	agcaacaggg	240
acaacgaggt	ggacttccaa	gagtaactgt	tcttctctgt	ctgcatcgcc	atgatgtgta	300
acgaattctt	tgaaggcttc	ccagataagc	agcccaggaa	gaaatgaaaa	ctcctctgat	360
gtggttgagg	ggtctgccag	ctggggccct	ccctgtcgcc	agtgggcact	tttttttttc	420
c						421

<210> 309

<211> 321

<212> DNA

<213> Homo sapien

<400> 309

accaaattgg	ggatgacgcc	ggtgcagcgg	gggggcccgg	gggccctggt	ggccctggga	60
tggggaaccg	cggtggtctc	cgcgagggtt	tgggcagtgg	catccggggc	cggggtcgcg	120
gccgtggacg	gggccggggc	cgaggccgcg	gagctcgcgg	aggcaaggcc	gaggataagg	180
agtggatgcc	cgtcaccaag	ttgggcccgt	tggtaagga	catgaagatc	aagtccttgg	240
aggagatcta	tctcttctcc	ctgcccatta	aggaatcaga	gatcattgat	ttcttctctg	300
ggcctctct	caaggatgag	g				321

<210> 310

<211> 381

<212> DNA

<213> Homo sapien

<400> 310

ttaaccagcc	atattggctc	aataaatagc	ttcggtaagg	agttaatttc	cttctagaaa	60
tcagtgccta	tttttcctgg	aaactcaatt	ttaaatagtc	caattccatc	tgaagccaag	120
ctgttgcat	tttcattcgg	tgacattctc	tcccatgaca	cccagaagg	gcagaagaac	180
cacatttttc	atttatagat	gtttgcatcc	tttgtattaa	aattattttg	aaggggttgc	240
ctcattggat	ggcttttttt	tttttcctcc	agggagaagg	ggagaaatgt	acttggaat	300
taatgtatgt	ttacatctct	ttgcaaattc	ctgtacatag	agatatattt	tttaagtgtg	360
aatgtaacaa	catactgtga	a				381

<210> 311

<211> 538

<212> DNA

<213> Homo sapien

<400> 311

tttgaattta	caccaagaac	ttctcaataa	aagaaaatca	tgaatgctcc	acaatttcaa	60
cataccacaa	gagaagttaa	tttcttaaca	ttgtgttcta	tgattatttg	taagaccttc	120
accaagtctt	gatatctttt	aaagacatag	ttcaaaattg	cttttgaaaa	tctgtattct	180
tgaaaatata	cttggttgtgt	attagggttt	taaaataccag	ctaaaggatt	acctcactga	240
gtcatcagta	ccctcctatt	cagctcccca	agatgatgtg	tttttgctta	ccctaagaga	300
ggttttcttc	ttatttttag	ataattcaag	tgcttagata	aattatgttt	tctttaagtg	360
tttatggtaa	actcttttaa	agaaaattta	atatgttata	gctgaatctt	tttggttaact	420
ttaaatcttt	atcatagact	ctgtacatat	gttcaaatta	gctgcttgcc	tgatgtgtgt	480
atcatcggtg	ggatgacaga	acaaacatat	ttatgatcat	gaataatgtg	ctttgttaa	538

<210> 312

<211> 176

<212> DNA

<213> Homo sapien

<400> 312

ggaggagcag	ctgagagata	gggtcagtga	atgcggttca	gcctgctacc	tctcctgtct	60
tcatagaacc	attgccttag	aattattgta	tgacacgttt	tttgttggtt	aagctgtaag	120
gttttgttct	ttgtgaacat	gggtattttg	aggggaggg	ggaggagta	gggaag	176

<210> 313

<211> 396

<212> DNA

<213> Homo sapien

<400> 313

ccagcacccc	caggccctgg	gggacctggg	ttctcagact	gccaaagaag	ccttgccatc	60
tggcgctccc	atggctcttg	caacatctcc	ccttcgtttt	tgaggggggc	atgccggggg	120
agccaccagc	ccctcactgg	gttcggagga	gagtcaggaa	gggccaagca	cgacaaagca	180
gaaacatcgg	atttggggaa	cgctgtgcaa	tcccttggtc	cgaggggctg	ggcgggagag	240
actgttctgt	tccttggtga	actgtgttgc	tgaaagacta	cctcgttctt	gtcttgatgt	300
gtcaccgggg	caactgcctg	ggggcgggga	tgggggcagg	gtggaagcgg	ctccccatth	360
tataccaaag	gtgtacatc	tatgtgatgg	gtgggg			396

<210> 314

<211> 311

<212> DNA

<213> Homo sapien

<400> 314

cctcaacatc	ctcagagagg	actggaagcc	agtccttacg	ataaaactcca	taattttatgg	60
cctgcagtat	ctctttcttg	agcccaaccc	cgaggaccca	ctgaacaagg	aggccgcaga	120
ggtcctgcag	aacaaccggc	ggctgtttga	gcagaacgtg	cagcgctcca	tgcggggtgg	180
ctacatcggc	tccacctact	ttgagcgctg	cctgaaatag	ggttggcgca	taccaccccc	240
cgccacggcc	acaagccctg	gcateccctg	caaatatatta	ttggggggcca	tgggtagggg	300
tttggggggc	g					311

<210> 315

<211> 336

<212> DNA

<213> Homo sapien

<400> 315

tttagaacat	ggttatcatc	caagactact	ctaccctgca	acattgaact	cccaagagca	60
aatccacatt	cctcttgagt	tctgcagctt	ctgtgtaa	agggcagctg	tcgtctatgc	120
cgtagaatca	catgatctga	ggaccattca	tggaaagctgc	taaatagcct	agtcctggga	180
gtcttccata	aagttttgca	tggagcaaac	aaacaggatt	aaactagggt	tggttccttc	240
agccctctaa	aagcataggg	cttagcctgc	aggcttcctt	gggctttctc	tgtgtgtgta	300
gttttgtaaa	cactatagca	tctgttaaga	tccagt			336

<210> 316

<211> 436

<212> DNA

<213> Homo sapien

<400> 316

aacatggtct	gcgtgcctta	agagagacgc	ttcctgcaga	acaggacctg	actacaaaga	60
atgtttccat	tggaattggt	ggtaaagact	tggagtttac	aatctatgat	gatgatgatg	120
tgtctccatt	cctggaaggt	cttgaagaaa	gaccacagag	aaaggcacag	cctgctcaac	180
ctgctgatga	acctgcagaa	aaggctgatg	aaccaatgga	acattaagtg	ataagccagt	240
ctatatatgt	attatcaa	atgtaagaat	acaggcacca	catactgatg	acaataatct	300
atactttgaa	ccaaaagttg	cagagtgggtg	gaatgctatg	tttttaggaat	cagtccagat	360
gtgagttttt	tccaagcaac	ctcactgaaa	cctatataat	ggaatacatt	tttctttgaa	420
agggctctgta	taatca					436

<210> 317

<211> 196

<212> DNA

<213> Homo sapien

<400> 317

tattccttgt	gaagatgata	tactatTTTT	gttaagcgtg	tctgtattta	tgtgtgagga	60
gctgctggct	tgcagtgcgc	gtgcacgtgg	agagctgggtg	cccgagagatt	ggacggcctg	120
atgctccctc	ccctgccctg	gtccagggaa	gctggccgag	ggtcctggct	cctgaggggc	180
atctgcccct	ccccca					196

<210> 318

<211> 381

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature
 <222> (1)...(381)
 <223> n = A,T,C or G

<400> 318

gacgcttnng	cogtaacgat	gacgggagac	atcctgctgt	tcgggacgtt	gctgatgaat	60
gccggggcgg	tgctgaactt	taagctgaaa	aagaaggaca	cncagggtt	tggggaggag	120
tncagggagc	ccaacacagg	tgacaacatc	cggaattct	tgctgancct	cagatacttt	180
cnaatcttca	tcnccctgtg	gaacatcttc	atgatgttct	gcatgattgt	gctgntcggc	240
tcttgaatcc	cancgatgaa	accannaact	cactttcccg	ggatgccgan	tctccattcc	300
tccattcctg	atgacttcaa	naatgttttt	gaccaaaaaa	ccgacaacct	tcccagaaag	360
tccaagctcg	tggtggngng	a				381

<210> 319
 <211> 506
 <212> DNA
 <213> Homo sapien

<400> 319

ctaagcttta	cgaatggggt	gacaacttat	gataaaaact	agagctagt	aattagccta	60
tttgtaaata	cctttgttat	aattgatagg	atacatcttg	gacatggaat	tgtaagcca	120
cctctgagca	gtgtatgtca	ggacttgctc	attagggttg	cagcagaggg	gcagaaggaa	180
ttatacaggt	agagatgtat	gcagatgtgt	ccatataatg	ccatatttac	attttgatag	240
ccattgatgt	atgcatctct	tggtgttact	ataagaacac	attaattcaa	tggaaataca	300
ctttgcta	attttaattg	tatagatctg	ctaataaatt	ctcttaaaaa	catactgtat	360
tctgttgctg	tgtgtttcat	tttaaattga	gcattaaggg	aatgcagcat	ttaaatcaga	420
actctgccaa	tgcttttata	tagaggcgtg	ttgccatttt	tgtcttatat	gaaatttctg	480
tccaagaaaa	ggcaggatta	catctt				506

<210> 320
 <211> 351
 <212> DNA
 <213> Homo sapien

<400> 320

ctgacctgca	ggacgaaacc	atgaagagcc	tgatccttct	tgccatcctg	gccgccttag	60
cggtagtaac	tttgtgttat	gaatcacatg	aaagcatgga	atcttatgaa	cttaatccct	120
tcattaacag	gagaaatgca	aataccttca	tatccctcca	gcagagatgg	agagctaaag	180
tccaagagag	gatccgagaa	cgctctaagc	ctgtccacga	gctcaatagg	gaagcctgtg	240
atgactacag	actttgcgaa	cgctacgcca	tggtttatgg	atacaatgct	gcctataatc	300
gctacttcag	gaagcgccga	gggaccaa	gagactgagg	gaagaaaaaa	a	351

<210> 321
 <211> 421
 <212> DNA
 <213> Homo sapien

<400> 321

ctcggaggcg	ttcagctgct	tcaagatgaa	gctgaacatc	tccttcccag	ccactggctg	60
ccagaaactc	attgaagtgg	acgatgaacg	caaacttcgt	actttctatg	agaagcgtat	120
ggccacagaa	gttgctgctg	acgctctggg	tgaagaatgg	aagggttatg	tggtccgaat	180
cagtgggtgg	aacgacaaac	aagggttccc	catgaagcag	ggtgtcttga	cccatggccg	240
tgccgcctg	ctactgagta	aggggcattc	ctgttacaga	ccaaggagaa	ctggagaaag	300
aaagagaaaa	tcagttcgtg	gttgcatgtg	ggatgcaaat	ctgagcggtc	tcaacttggt	360

tattgtaaaa aaaggagaga aggatatcc tggactgact gatactacag tgccctgccg 420
c 421

<210> 322
<211> 521
<212> DNA
<213> Homo sapien

<400> 322

agcagctctc	ctgccacagc	tectcacccc	ctgaaaatgt	tcgcctgctc	caagtttgtc	60
tccactccct	ccttggtcaa	gagcacctca	cagctgctga	gccgtccgct	atctgcagtg	120
gtgctgaaac	gaccggagat	actgacagat	gagagcctca	gcagcttggc	agtctcatgt	180
ccccttacct	cacttggtctc	tagccgcagc	ttccaaacca	gcgccatttc	aagggacatc	240
gacacagcag	ccaagttcat	tggagctggg	gctgccacag	ttgggggtggc	tggtttctggg	300
gctgggattg	gaactgtgtt	tgggagcctc	atcattgggt	atgccaggaa	ccctttctctg	360
aagcaacagc	tctttctcta	cgccattctg	ggctttgccc	tctcgagggc	catggggctc	420
ttttgtctga	tggtagcctt	tctcatcctc	tttgccatgt	gaaggagccg	tctccacctc	480
ccatagtctt	cccgctctg	gttggccccg	tgtgttctt	t		521

<210> 323
<211> 435
<212> DNA
<213> Homo sapien

<400> 323

ccgaggtcgc	acgcgtgaga	cttctccgcc	gcagacgccg	ccgcgatgcg	ctacgtcgcc	60
tcctacctgc	tggctgccct	agggggcaac	tcctcccca	gcgccaagga	catcaagaag	120
atcttggaac	gcgtgggtat	cgaggcggac	gacgaccggc	tcaacaaggt	tatcagttag	180
ctgaatggaa	aaaacattga	agacgtcatt	gcccagggtg	ttggcaagct	tgccagtgtg	240
cctgctgggt	gggtgttagc	cgtctctgct	gccccaggct	ctgcagcccc	tgctgctggt	300
tctgcccctg	ctgcagcaga	ggagaagaaa	gatgagaaga	aggaggagtc	tgaagagtca	360
gatgatgaca	tgggatttgg	cctttttgat	taaattcctg	ctcccctgca	aataaagcct	420
ttttacacat	ctcaa					435

<210> 324
<211> 521
<212> DNA
<213> Homo sapien

<400> 324

aggagatcga	ctttcgggtgc	ccgcaagacc	agggctggaa	cgccgagatc	acgtgcaga	60
tgggtgcagta	caagaatcgt	caggccatcc	tggcgggtcaa	atccacgcgg	cagaagcagc	120
agcacctggt	ccagcagcag	ccccctcgc	agccgcagcc	gcagccgcag	ctccagcccc	180
aaccccagcc	tcagcctcag	ccgcaacccc	agccccaatc	acaaccccag	cctcagcccc	240
aacccaagcc	tcagccccag	cagctccacc	cgtatccgca	tccacatcca	catccacact	300
ctcatcctca	ctgcaccca	caccctcacc	cgcacccgca	tccgcaccaa	ataccgcacc	360
cacacccaca	gcgcactcgc	cagccgcacg	ggcaccggct	tctccgcagc	acctccaact	420
ctgcctgaaa	ggggcagctc	ccgggcaaga	caaggttttg	aggacttgag	gaagtgggac	480
gagcacatct	ctattgtctt	cacttggtgc	aaaagcaaaa	c		521

<210> 325
<211> 451
<212> DNA
<213> Homo sapien

<400> 325

```

attttcattt ccattaacct ggaagctttc atgaatattc tcttctttta aaacatttta      60
acattatttta aacagaaaaa gatgggctct ttctgggttag ttgttacatg atagcagaga      120
tatttttact tagattactt tgggaatgag agattgttgt cttgaactct ggcactgtac      180
agtgaatgtg tctgtagttg tgttagtttg cattaagcat gtataacatt caagtatgtc      240
atccaaataa gaggcataata cattgaattg tttttaatcc tctgacaagt tgactcttcg      300
acccccaccc ccacccaaga cattttaata gtaaataagag agagagagaa gagttaatga      360
acatgaggta gtgttccact ggcaggatga cttttcaata gctcaaata atttcagtgc      420
ctttatcact tgaattatta acttaatttg a                                     451

```

<210> 326

<211> 421

<212> DNA

<213> Homo sapien

<220>

<221> misc_feature

<222> (1)...(421)

<223> n = A,T,C or G

<400> 326

```

cgcggtcgta agggctgagg atttttggtc cgcacgctcc tgctcctgac tcaccgctgt      60
tcgctctcgc cgaggaacaa gtgggtcagg aagccgcgcg gcaacagcca tggcttttaa      120
ggataccgga aaaacacccg tggagccgga ggtggcaatt caccgaattc gaatcacccct      180
aacaagccgc aacgtaaaat ccttggaaaa ggtgtgtgct gacttgataa gaggcgcaaa      240
agaaaagaat ctcaaagtga aaggaccagt tcgaatgcct accaagactt tgagantcac      300
tacaagaaaa actccttgtg gtgaagggtc taagacgtgg gatcgtttcc agatgagaat      360
tcacaagcga ctcatgtact tgcacagtcc ttctgagatt gttaagcaga ttacttccat      420
c                                     421

```

<210> 327

<211> 456

<212> DNA

<213> Homo sapien

<400> 327

```

atcttgacga ggctgcggtg tctgctgcta ttctccgagc ttcgcaatgc cgcctaagga      60
cgacaagaag aagaaggacg ctggaaagtc ggccaagaaa gacaaagacc cagtgaacaa      120
atccgggggc aaggccaaaa agaagaagtg gtccaaaggc aaagttcggg acaagctcaa      180
taacttagtc ttgtttgaca aagctaccta tgataaactc tgtaaggaag ttcccaacta      240
taaacttata accccagctg tggctctctg gagactgaag attcgaggct ccctggccag      300
ggcagccctt caggagctcc ttagtaaagg acttatcaaa ctggtttcaa agcacagagc      360
tcaagtaatt tacaccagaa ataccaaggg tggagatgct ccagctgctg gtgaagatgc      420
atgaataggt ccaaccagct gtacatttgg aaaaat                                     456

```

<210> 328

<211> 471

<212> DNA

<213> Homo sapien

<400> 328

```

gtggaagtga catcgtcttt aaaccctgcg tggcaatccc tgacgcaccg ccgtgatgcc      60
caggaagac agggcgacct ggaagtccaa ctacttcctt aagatcatcc aactattgga      120

```

```

tgattatccg aaatgtttca ttgtgggagc agacaatgtg ggctccaagc agatgcagca 180
gatccgcatg tcccttcgcg ggaaggctgt ggtgctgatg ggcaagaaca ccatgatgcg 240
caaggccatc cgagggcacc tggaaaacaa cccagctctg gagaaactgc tgcctcatat 300
ccgggggaat gtgggctttg tgttcaccaa ggaggacctc actgagatca gggacatgtt 360
gctggccaat aagggtgccag ctgctgcccg tgctgggtgc attgccccat gtgaagtcac 420
tgtgccagcc cagaacactg gtctcggggc cgagaagacc tcctttttcc a 471

```

```

<210> 329
<211> 278
<212> DNA
<213> Homo sapien

```

```

<220>
<221> misc_feature
<222> (1)...(278)
<223> n = A,T,C or G

```

```

<400> 329
gtttaaactt aagcttggtg ccgagctcgg atccactagt ccagtgtggt ggaattctag 60
aaattgagat gcccccccag gccagcaaat gttccttttt gttcaaagtc tatttttatt 120
ccttgatatt tttctttttt tttttttttt ttgnggatgg ggacttgtga atttttctaa 180
aggtgctatt taacatggga gganagcgtg tgcggctcca gccagccccg ctgctcactt 240
tccacctctt ctccacctgc ctctggcttc tcaggcct 278

```

```

<210> 330
<211> 338
<212> DNA
<213> Homo sapien

```

```

<400> 330
ctcaggett c aacatogaat acgcgcgagg ccccttcgcc ctattcttca tagccgaata 60
cacaacatt attataataa acaccctcac cactacaatc ttcctaggaa caacatatga 120
cgcactctcc cctgaactct acacaacata tttgtccacc aagaccctac ttctaacctc 180
cctgttctta tgaattcgaa cagcataccc ccgattccgc tacgaccaac tcataacct 240
cctatgaaaa aacttcttac cactcaccct agcattactt atatgatatg tctccatacc 300
cattacaate tccagcattc cccctcaaac ctaaaaaa 338

```

```

<210> 331
<211> 2820
<212> DNA
<213> Homo sapiens

```

```

<400> 331
tggcaaaatc ctggagccag aagaaaggac agcagcattg atcaatctta cagctaacat 60
gttgtaacct gaaaacaatg cccagactca atttagtgag ccacagtaca cgaacctggg 120
gtcctgaac agcatggacc agcagattcg gaacggctcc tcgtccacca gtccctataa 180
cacagaccac ggcgagaaca gcgtcacggc gccctcgccc tacgcacagc ccagccccac 240
cttogatgct ctctctccat caccgcccac cccctccaac accgactacc caggccccga 300
cagttccgac gtgtccttcc agcagtcgag caccgccaag tcggccacct ggacgtattc 360
cactgaactg aagaaactct actgccaaat tgcaaagaca tgccccatcc agatcaaggt 420
gatgacccca cctcctcagg gagctgttat ccgcgccatg cctgtctaca aaaaagctga 480
gcacgtcacg gaggtggtga agcgggtgcc caaccatgag ctgagccgtg agttcaacga 540
gggacagatt gccctccta gtcatttgat tcgagtagag gggaacagcc atgccagta 600

```

```

tgtagaagat cccatcacag gaagacagag tgtgctggta ccttatgagc caccgccaggt 660
tggcactgaa ttcacgacag tcttgtagaa tttcatgtgt aacagcagtt gtgttgagg 720
gatgaaccgc cgtccaattt taatcattgt tactctggaa accagagatg ggcaagtcct 780
gggccgacgc tgctttgagg cccggatctg tgcttgccca ggaagagaca ggaaggcgga 840
tgaagatagc atcagaaagc agcaagtttc ggacagtaca aagaacggtg atggtacgaa 900
gcgcccgttt cgtcagaaca cacatgggat ccagatgaca tccatcaaga aacgaagatc 960
cccagatgat gaactgttat acttaccagt gaggggcccgt gagacttatg aaatgctgtt 1020
gaagatcaaa gagtccctgg aactcatgca gtaccttctt cagcacacaa ttgaaacgta 1080
caggcaacag caacagcagc agcaccagca cttacttcag aaacagacct caatacagtc 1140
tccatcttca tatggtaaca gctccccacc tctgaacaaa atgaacagca tgaacaagct 1200
gccttctgtg agccagctta tcaacctca gcagcgcaac gccctcactc ctacaaccat 1260
tctgatggc atgggagcca acattcccat gatgggcacc cacatgccaa tggctggaga 1320
catgaatgga ctacgcccc cccaggcact cctccccca ctctccatgc catccacctc 1380
ccactgcaca cccccacctc cgtatcccac agattgcagc attgtcagtt tcttagcgag 1440
gttgggctgt tcatcatgtc tggactattt cagcaccagc gggctgacca ccatctatca 1500
gattgagcat tactccatgg atgatctggc aagtctgaaa atccctgagc aatttcgaca 1560
tgcatctgg aagggcatcc tggaccaccg gcagctccac gaattctcct ccccttctca 1620
tctcctgcgg accccaagca gtgctctac agtcagtgtg ggctccagtg agaccgggg 1680
tgagcgtgtt attgatgctg tgcgattcac cctccgccag accatctctt tcccaccccg 1740
agatgagtg aatgacttca actttgacat gtagctcttc ctatccctct cctaactgcc 1800
caaagaggag ggggagttag cctcaccatg tgagctcttc ctatccctct cctaactgcc 1860
agccccctaa aagcactcct gcttaatctt caaagccttc tccctagctc ctccccttcc 1920
tcttgcttga tttcttaggg gaaggagaag taaggaggta cctcttacct aacatctgac 1980
ctggcatcta attctgattc tggttttaag ccttcaaaac tatagcttgc agaactgtag 2040
ctgccatggc taggtagaag tgagcaaaaa agagttgggt gtctccttaa gctgcagaga 2100
tttctcattg acttttataa agcatgttca cctttatagt ctaagactat atatataaat 2160
gtataaatat acagtataga tttttgggtg gggggcattg agtattgttt aaaatgtaat 2220
ttaaatgaaa gaaaattgag ttgcacttat tgaccatttt ttaatttact tgttttggat 2280
ggcttgtcta tactccttcc ctttaaggggt atcatgtatg gtgataggta tctagagctt 2340
aatgctacat gtgagtgcga tgatgtacag attcctttcag ttctttggat tctaaataca 2400
tgccacatca aacctttgag tagatccatt tccattgctt attatgtagg taagactgta 2460
gatatgtatt cttttctcag tgttgggtata ttttatatta ctgacatttc ttctagtgat 2520
gatgggtcac gttggggtga tttaatccag ttataagaag aagttcatgt ccaaacggtc 2580
ctctttagtt tttggttggg aatgaggaaa attcttaaaa ggcccatagc agccagttca 2640
aaaaaccccg acgtcatgta tttgagcata tcagtaaccc ccttaaattt aataccaga 2700
taccttactc tacaatgttg attgggaaaa catttgctgc ccattacaga ggtattaaaa 2760
ctaaatttca ctactagatt gactaactca aatacacatt tgctactgtt gtaagaattc 2820

```

<210> 332

<211> 2270

<212> DNA

<213> Homo sapiens

<400> 332

```

tcgttgatat caaagacagt tgaaggaaat gaattttgaa acttcacggt gtgccaccct 60
acagtactgc cctgaccctt acatccagcg tttcgtagaa acccagctca tttctcttgg 120
aaagaaagtt attaccgata caccatgtcc cagagcacac agacaaatga attcctcagt 180
ccagaggttt tccagcatat ctgggatttt ctggaacagc ctatatgttc agttcagccc 240
attgacttga actttgtgga tgaaccatca gaagatgggt cgacaaacaa gattgagatt 300
agcatggact gtatccgcat gcaggactcg gacctgagtg accccatgtg gccacagtac 360
acgaacctgg ggctcctgaa cagcatggac cagcagattc agaacggctc ctcgccacc 420
agtccttata acacagacca cgcgcagaac agcgtcacgg cgccctcgcc ctacgcacag 480
cccagctcca ccttcgatgc tctctctcca tcaccgccca tcccctccaa caccgactac 540

```

```

ccaggccccgc acagttttcga cgtgtccttc cagcagtcga gcaccgcca gtcggccacc 600
tggaactgatt ccactgaact gaagaaactc tactgcaaaa ttgcaaagac atgccccatc 660
cagatcaagg tgatgacccc acctcctcag ggagctgtta tccgcgccat gcctgtctac 720
aaaaaagctg agcacgtcac ggaggtggtg aagcgggtgcc ccaacatga gctgagccgt 780
gaattcaacg agggacagat tgcctcctcct agtcatttga ttcgagtaga ggggaacagc 840
catgcccagt atgtagaaga tcccatcaca ggaagacaga gtgtgctggt accttatgag 900
ccaccccagg ttggcaactga attcacgaca gtcttgtaca atttcattgt taacagcagt 960
tgtgttgag ggatgaaccg ccgtccaatt ttaatcattg ttactctgga aaccagagat 1020
gggcaagtcc tgggcccagc ctgctttgag gcccgatct gtgcttgccc aggaagagac 1080
aggaaggcgg atgaagatag catcagaaag cagcaagttt cggacagtac aaagaacggg 1140
gatggtacga agcgcocgtt tcgtcagaac acacatggta tccagatgac atccatcaag 1200
aaacgaagat cccagatga tgaactgtta tacttaccag tgaggggccc tgagacttat 1260
gaaatgctgt tgaagatcaa agagtccctg gaactcatgc agtaccttc tcagcacaca 1320
attgaaacgt acaggcaaca gcaacagcag cagcaccagc acttacttca gaaacagacc 1380
tcaatacagt ctccatcttc atatggtaac agctccccac ctctgaacaa aatgaacagc 1440
atgaacaagc tgctttctgt gagccagctt atcaacctc agcagcgcaa cgcctcact 1500
cctacaacca ttctgatgg catgggagcc aacattccca tgatgggcac ccacatgcca 1560
atggctggag acatgaatgg actcagcccc acccaggcac tccctcccc actctccatg 1620
ccatccacct cccactgcac acccccacct ccgtatccaa cagattgcag cattgtcggg 1680
ttcttagcga ggttgggctg ttcactatgt ctggactatt tcaagacca ggggctgacc 1740
accatctatc agattgagca ttactccatg gatgatctgg caagtctgaa aatccctgag 1800
caatttcgac atgcgatctg gaagggcatc ctggaccacc ggcagctcca cgaattctcc 1860
tccccttctc atctcctggg gaccccaagc agtgccctca cagtcagtgt gggctccagt 1920
gagacccggg gtgagcgtgt tattgatgct gtgcgattca cctccgcca gaccatctct 1980
ttcccacccc gagatgagtg gaatgacttc aactttgaca tggatgctcg ccgaataag 2040
caacagcgca tcaaagagga gggggagtga gcctcaccat gtgagctctt cctatccctc 2100
tcctaactgc cagcccccta aaagcactcc tgcttaatct tcaaagcctt ctccctagct 2160
cctcccttcc ctcttgtctg atttcttagg ggaaggagaa gtaagaggct acctcttacc 2220
taacatctga cctggcatct aattctgatt ctggctttaa gccttcaaaa 2270

```

<210> 333

<211> 2816

<212> DNA

<213> Homo sapiens

<400> 333

```

tcgttgatat caaagacagt tgaaggaaat gaattttgaa acttcacggg gtgccaccct 60
acagtactgc cctgaccctt acatccagcg ttctgtagaa acccagctca tttctcttgg 120
aaagaaagtt attaccgata caccatgtcc cagagcacac agacaaatga attcctcagt 180
ccagaggttt tccagcatat ctgggatttt ctggaacagc ctatatgttc agttcagccc 240
attgacttga acttttgtga tgaacctca gaagatggtg cgacaaaca gattgagatt 300
agcatggact gtatccgcat gcaggactcg gacctgagtg acccatgtg gccacagtac 360
acgaacctgg ggctcctgaa cagcatggac cagcagattc agaacggctc ctctgccacc 420
agtccctata acacagacca cgcgcagaac agcgtcacgg cgcctcgcgc ctacgcacag 480
cccagctcca ccttcgatgc tctctctcca tcaccgcca tccctccaa caccgactac 540
ccaggccccgc acagttttcga cgtgtccttc cagcagtcga gcaccgcca gtcggccacc 600
tggaactgatt ccactgaact gaagaaactc tactgcaaaa ttgcaaagac atgccccatc 660
cagatcaagg tgatgacccc acctcctcag ggagctgtta tccgcgccat gcctgtctac 720
aaaaaagctg agcacgtcac ggaggtggtg aagcgggtgcc ccaacatga gctgagccgt 780
gaattcaacg agggacagat tgcctcctcct agtcatttga ttcgagtaga ggggaacagc 840
catgcccagt atgtagaaga tcccatcaca ggaagacaga gtgtgctggt accttatgag 900
ccaccccagg ttggcaactga attcacgaca gtcttgtaca atttcattgt taacagcagt 960
tgtgttgag ggatgaaccg ccgtccaatt ttaatcattg ttactctgga aaccagagat 1020

```

```

gggcaagtcc tggggccgacg ctgcttttgag gcccggatct gtgcttgccc aggaagagac 1080
aggaaggcgg atgaagatag catcagaaaag cagcaagttt cggacagtac aaagaacggg 1140
gatggtacga agcgcgccgtt tcgtcagaac acacatggta tccagatgac atccatcaag 1200
aaacgaagat ccccgatga tgaactgtta tacttaccag tgaggggccc tgagacttat 1260
gaaatgctgt tgaagatcaa agagtccctg gaactcatgc agtaccttcc tcagcacaca 1320
attgaaacgt acaggcaaca gcaacagcag cagcaccagc acttacttca gaaacatctc 1380
ctttcagcct gcttcaggaa tgagcttggt gagccccgga gagaaactcc aaaacaatct 1440
gacgtcttct ttagacattc caagccccca aaccgatcag tgtaccata gagccctatc 1500
tctatatatt aagtgtgtgt gttgtatttc catgtgtata tgtgagtgtg tgtgtgtgta 1560
tgtgtgtgcg tgtgtatcta gccctcataa acaggacttg aagacacttt ggctcagaga 1620
cccaactgct caaaggcaca aagccactag tgagagaatc ttttgaaggg actcaaacct 1680
ttacaagaaa ggatgttttc tgcagatttt gtatccttag accggccatt ggtgggtgag 1740
gaaccactgt gtttgtctgt gagctttctg ttgtttcctg ggagggaggg gtcaggtggg 1800
gaaaggggca ttaagatgtt tattggaacc cttttctgtc ttcttctgtt gtttttctaa 1860
aattcacagg gaagcttttg agcaggtctc aaacttaaga tgtcttttta agaaaaggag 1920
aaaaaagttg ttattgtctg tgcataagta agttgtaggt gactgagaga ctcagtcaga 1980
cccttttaat gctggtcatt taataatatt gcaagtagta agaaacgaag gtgtcaagtg 2040
tactgctggg cagcgagggt atcattacca aaagtaatca actttgtggg tggagagttc 2100
tttgtgagaa cttgcattat ttgtgtcctc cctcatgtg taggtagaac atttcttaat 2160
gctgtgtacc tgcctctgcc actgtatgtt ggcactctgt atgctaaagt ttttcttgta 2220
catgaaaccc tggagacct actacaaaaa aactgttggt tggcccccac agcaggtgaa 2280
ctcattttgt gcttttaata gaaagacaaa tccaccccag taatattgcc cttacgtagt 2340
tgtttaccat tattcaaagc tcaaaataga atttgaagcc ctctcacaaa atctgtgatt 2400
aatttgctta attagagctt ctatccctca agcctaccta ccataaaacc agccatatta 2460
ctgatactgt tcagtgcatt tagccaggag acttacgttt tgagtaagtg agatccaagc 2520
agacgtgtta aaatcagcac tcttgactg gaaattaaag attgaaaggg tagactactt 2580
ttcttttttt tactcaaaag tttagagaat ctctgtttct ttccatttta aaaacatatt 2640
ttaagataat agcataaaga ctttaaaaaat gttcctcccc tccatcttcc cacaccagc 2700
caccagcact gtattttctg tcaccaagac aatgatttct tgttattgag gctgttgctt 2760
ttgtggatgt gtgattttta ttttcaataa acttttgcac ctgggtttaa aagaaa 2816

```

<210> 334

<211> 2082

<212> DNA

<213> Homo sapiens

<400> 334

```

agatgctaca ggaactgcac acccaggctg tatgatacag cctattgctc ccgggctgca 60
aacctgtcca gcatgtgatg tgggtgggata ctgaattgaa taccgaatac tgtaggcaat 120
tgtaacacag tggtaagtct ttgtgtatct aaacatagct aaacaccaaa aggtatagta 180
agaatatggg attataatct tatggaacta tcattgtata tgtggtttgt caaccagaat 240
gtagttatac agcacaggac tgtgcttatg atgtgccaag cacagctctc agtactaact 300
cctttaatct tcatatcaac cctaggaggt aacttcttaa gtagattcat attgtaaggg 360
tctcgggggtg gggggggttg caaaatcctg gagccagaag aaaggacagc agcattgatc 420
aatcttacag ctaacatggt gtacctggaa aacaatgccc agactcaatt tagtgagcca 480
cagtacacga acctggggct cctgaacagc atggaccagc agattcagaa cggctcctcg 540
tccaccagtc cctataacac agaccacgag cagaacagcg tcacggcgcc ctgcacctac 600
gcacagccca gctccacctt cgatgctctc tctccatcac ccgccatccc ctccaacacc 660
gactaccag gccgcacag tttcgacgtg tcttccagc agtcgagcac cgccaagtgc 720
gocacctgga cgtattccac tgaactgaag aaactctact gccaaattgc aaagacatgc 780
cccatccaga tcaagggtgat gacccacct cctcaggag ctgttatccg cgccatgctt 840
gtctacaaaa aagctgagca cgtcacggag gtggtgaagc ggtgccccaa ccatgagctg 900
agccgtgaat tcaacgaggg acagattgcc cctcctagtc atttgattcg agtagaggg 960

```


aacagccatg	cccagtatgt	agaagatccc	atcacaggaa	gacagagtgt	gctgggtacct	1020
tatgagccac	cccaggttgg	caactgaattc	acgacagtct	tgtacaattt	catgtgtaac	1080
agcagttgtg	ttggagggat	gaaccgccgt	ccaattttta	tcattgttac	tctggaaacc	1140
agagatgggc	aagtcctggg	ccgacgctgc	tttgaggccc	ggatctgtgc	ttgccagga	1200
agagacagga	aggcggatga	agatagcatc	agaaagcagc	aagtttcgga	cagtacaaag	1260
aacggtgatg	gtacgaagcg	cccgtctcgt	cagaacacac	atggtatcca	gatgacatcc	1320
atcaagaaac	gaagatcccc	agatgatgaa	ctgttatact	taccagtgag	gggccgtgag	1380
acttatgaaa	tgctgttgaa	gatcaaagag	tccttgaac	tcatgcagta	ccttccctcag	1440
cacacaattg	aaacgtacag	gcaacagcaa	cagcagcagc	accagcactt	acttcagaaa	1500
cagtgaagtgt	atcaacgtgt	catttttagga	ggcatgagtgt	acggtgactt	tatttggtac	1560
agcaatagg	tgattgatga	gcaatgtgga	acataatggg	agatagcaga	ttgtcataga	1620
ttcagatgac	ctgggtatggc	aaccctcttt	cagttgcaac	cttttttacg	tgtcttatta	1680
taaccttccc	ttcagaattc	cacttatgtt	ctgaaattaa	atacaaaacca	tttctggtga	1740
attacaaaga	aactcacact	aacagttctc	ttctctatat	gcctggtcca	tacacactaa	1800
cagtaagtac	acactctatt	tggtagtgat	gtgtatattt	gaaaacatga	aatcttttct	1860
catcccaatg	gattgtctta	taaatctcct	gggatgcaca	ctatccactt	ttgggaataa	1920
cactgtagac	cagggatagc	aaataggctt	tactataata	taaagtgact	tgtttgaatg	1980
ctgtaatgag	aagaattctg	agacctagtgt	catgataatt	ggggaatat	ctgggtgcag	2040
aaggataagg	tagcatcatg	ttgccgtatt	ttagcatctc	tg		2082

<210> 335

<211> 4849

<212> DNA

<213> Homo sapiens

<400> 335

cgttgataac	aaagacagtt	gaaggaaatg	aattttgaaa	cttcacggtg	tgccacccta	60
cagtactgcc	ctgaccctta	catccagcgt	ttcgtagaaa	ccccagctca	tttctcttgg	120
aaagaaagt	attaccgatc	caccatgtcc	cagagcacac	agacaaatga	attcctcagt	180
ccagaggttt	tccagcatat	ctgggatttt	ctggaacagc	ctatatgttc	agttcagccc	240
attgacttga	actttgtgga	tgaaccatca	gaagatgggt	cgacaaacaa	gattgagatt	300
agcatggact	gtatccgcat	gcaggactcg	gacctgagt	accccatgtg	gccacagtac	360
acgaacctgg	ggctcctgaa	cagcatggac	cagcagattc	agaacggctc	ctcgtccacc	420
agtccttata	acacagacca	cgcgcagaa	agcgtcacgg	cgccctcgcc	ctacgcacag	480
cccagctcca	ccttcgatgc	tctctctcca	tcacccgcca	tcccctccaa	caccgactac	540
ccaggcccg	acagtttctga	cgtgtccttc	cagcagtcga	gcaccgcaa	gtcggccacc	600
tgagcgtatt	ccactgaact	gaagaaactc	tactgccaaa	ttgcaaagac	atgccccatc	660
cagatcaagg	tgatgacccc	acctcctcag	ggagctgtta	tccgcgccat	gcctgtctac	720
aaaaaagctg	agcacgtcac	ggaggtgggt	aagcgggtgc	ccaacatga	gctgagccgt	780
gaattcaacg	aggacagat	tgcctctcct	agtcatttga	ttcgagtaga	ggggaacagc	840
catgccagct	atgtagaaga	tcccatcaca	ggaagacaga	gtgtgctggt	accttatgag	900
ccaccccagg	ttggcactga	attcacgaca	gtcttgtaga	atttcattgt	taacagcagt	960
tgtgttgag	ggatgaaccg	cgttccaatt	ttaatcattg	ttactctgga	aaccagagat	1020
gggcaagtcc	tgggcccagc	ctgctttgag	gcccggatct	gtgcttgccc	aggaagagac	1080
aggaaggcgg	atgaagatag	catcagaaag	cagcaagttt	cggacagtac	aaagaacggt	1140
gatggtacga	agcgcctcgt	tgcgcagaac	acacatggta	tccagatgac	atccatcaag	1200
aaacgaagat	cccagatga	tgaactgtta	tacttaccag	tgaggggccc	tgagacttat	1260
gaaatgctgt	tgaagatcaa	agagtcctct	gaactcatgc	agtaccttcc	tcagcacaca	1320
attgaaacgt	acaggcaaca	gcaacagcag	cagcaccagc	acttacttca	gaaacagacc	1380
tcaatacagt	ctccatcttc	atatggtaac	agctccccac	ctctgaacaa	aatgaacagc	1440
atgaacaagc	tgccttctgt	gagccagctt	atcaaccttc	agcagcgcaa	cgccctcact	1500
cctacaacca	ttcctgatgg	catgggagcc	aacattccca	tgatgggcac	ccacatgcca	1560
atggctggag	acatgaatgg	actcagcccc	accagggcac	tccctcccc	actctccatg	1620

ccatccacct cccagtgcac acccccacct ccgtatccca cagattgcag cattgtcagt 1680
ttcttagcga ggttgggctg ttcatcatgt ctggactatt tcacgaccca ggggctgacc 1740
accatctatc agattgagca ttactccatg gatgatctgg caagtctgaa aatccctgag 1800
caatttcgac atgcgatctg gaagggcatc ctggaccacc ggcagctcca cgaattctcc 1860
tccccctctc atctcctgcg gacccccagc agtgccctca cagtcagtgt gggctccagt 1920
gagacccggg gtgagcgtgt tattgatgct gtgcgattca cctccgcca gaccatctct 1980
ttcccacccc gagatgagtg gaatgacttc aactttgaca tggatgctcg ccgcaataag 2040
caacagcgca tcaaagagga gggggagtga gcctcaccat gtgagctctt cctatccctc 2100
tcctaaactgc cagcycccta aaagcactcc tgcttaatct tcaaagcctt ctccctagct 2160
cctccctctc ctcttgtctg atttcttagg ggaaggagaa gtaagaggct acctcttacc 2220
taacatctga cctggcatct aattctgatt ctggctttaa gccttcaaaa ctatagcttg 2280
cagaactgta gctgccatgg ctaggtagaa gtgagcaaaa aagagtggg tgtctcctta 2340
agctgcagag atttctcatt gacttttata aagcatgttc acccttatag tctaagacta 2400
tatatatataa tgtataaata tacagtatag atttttgggt ggggggcatt gagtattgtt 2460
taaaatgtaa tttaaatgaa agaaaattga gttgcactta ttgaccattt tttaatctac 2520
ttgttttgga tggcttgtct atactccttc ccttaagggg tatcatgtat ggtgataggt 2580
atctagagct taatgctaca tgtgagtgc gatgatgtac agattctttc agttctttgg 2640
attctaaata catgccacat caaacctttg agtagatcca ttccattgc ttattatgta 2700
ggtaagactg tagatatgta ttcttttctc agtgttggtat tattttatat tactgacatt 2760
tcttctagtgc atgatggctc acgttggggg gatttaatcc agttataaga agaagttcat 2820
gtccaaacgt cctcttttagt ttttgggttg gaatgaggaa aattcttaaa aggccctag 2880
cagccagttc aaaaacaccc gacgtcatgt atttgagcat atcagtaacc cctttaaatt 2940
taataccaga taccttatct tacaatattg attgggaaaa catttgcctg cattacagag 3000
gtattaaaac taaatttcac tactagattg actaactcaa atacacattt gctactgttg 3060
taagaattct gattgatttg attgggatga atgccatcta tctagtctta acagtgaagt 3120
tttactgtct attaatattc agggtaataa ggaatcattc agaaatgttg agtctgtact 3180
aaacagtaag atatctcaat gaaccataaa ttcaactttg taaaaatctt ttgaagcata 3240
gataaatattg tttggtaaat gtttcttttg tttggtaaat gtttctttta aagaccctcc 3300
tattctataa aactctgcat gtagaggctt gtttaccttt ctctctctaa gggtttacaat 3360
aggagtgggtg atttgaaaaa tataaaatta tgagattgggt tttcctgtgg cataaattgc 3420
atcactgtat cattttcttt tttaacgggt aagagtttca gtttgttgga aagtaactgt 3480
gagaacccag tttcccgtec atctccctta gggactaccc atagacatga aaggteccca 3540
cagagcaaga gataagtctt tcatggctgc tgttgcttaa accacttaaa cgaagagttc 3600
ccttgaaact ttgggaaaac atgttaatga caatattcca gatctttcag aaatataaca 3660
catttttttg catgcattga aatgagctct gaaatcttcc catgcattct ggtcaagggc 3720
tgtcatttgc cataagcttc cattttaatt ttaaatgtca aaagggccag cgtggctcta 3780
aaaggtaatg tgtggattgc ctctgaaaag tttgtatata ttttgtgtga aattgcatac 3840
tttgtatttt gattattttt tttttcttct tgggatagtg ggatttccag aaccacactt 3900
gaaacctttt tttatogttt ttgtattttc atgaaaatac catttagtaa gaataccaca 3960
tcaaataaga aataatgcta caattttaag aggggagggg agggaaaagt tttttttatt 4020
atttttttta aattttgtat gttaaagaga atgagtcctt gatttcaaag ttttgttgta 4080
cttaaatggt aataagcact gtaaacttct gcaacaagca tgcagctttg caaaccctatt 4140
aaggggaaga atgaaagctg ttcccttggtc ctagttaagaa gacaaaactgc tcccttact 4200
ttgctgaggg tttgaataaa cctaggactt ccgagctatg tcagtactat tcaggtaaca 4260
ctagggcctt ggaaattcct gtactgtgtc tcatggattt ggcactagcc aaagcgaggc 4320
acccttactg gcttacctcc tcatggcagc ctactctcct tgagtgtatg agtagccagg 4380
gtaaggggta aaaggatagt aagcatagaa accactagaa agtgggctta atggagtct 4440
tgtggcctca gctcaatgca gttagctgaa gaattgaaaa gtttttggtt ggagacgttt 4500
ataaacagaa atggaaagca gagttttcat taaatccttt tacctttttt ttttcttggt 4560
aatcccttaa aataacagta tgtgggatat tgaatgttaa agggatattt tttttctatt 4620
atttttataa ttgtacaaaa ttaagcaaat gttaaaagtt ttatatgctt tattaatgtt 4680
ttcaaaaagg attatacatg tgatacatct ttttaagcttc agttgcttgt cttctgggtac 4740
tttctgttat gggcttttgg ggagccagaa gccaatctac aatctctttt tgtttgccag 4800
gacatgcaat aaaattttaa aaataaataa aaactaatta agaaataaa 4849

<210> 336
 <211> 1386
 <212> DNA
 <213> Homo sapiens

<400> 336
 atgtttgtacc tggaaaacaa tgcccagact caatttagtg agccacagta cacgaacctg 60
 gggctcctga acagcatgga ccagcagatt cagaacggct cctcgtccac cagtccctat 120
 aacacagacc acgcgcagaa cagcgtcacg gcgcctcgc cctacgcaca gccagctcc 180
 accttcgatg ctctctctcc atcaccgcgc atcccctcca acaccgacta cccaggcccg 240
 cacagtttcg acgtgtcctt ccagcagtcg agcaccgcca agtcggccac ctggacgtat 300
 tccactgaac tgaagaaact ctactgccaa attgcaaaga catgccccat ccagatcaag 360
 gtgatgaccc cacctcctca gggagctgtt atccgcgcca tgcctgtcta caaaaaagct 420
 gagcacgtca cggaggtggt gaagcgggtgc cccaaccatg agctgagccg tgaattcaac 480
 gagggacaga ttgcccctcc tagtcatttg attcgagtag aggggaacag ccatgcccag 540
 tatgtagaag atcccatcac aggaagacag agtgtgctgg taccttatga gccaccccag 600
 gttggcactg aattcacgac agtcttgtac aatttcatgt gtaacagcag ttgtgttga 660
 gggatgaacc gccgtccaat tttaatcatt gttactctgg aaaccagaga tgggcaagtc 720
 ctgggccgac gctgctttga ggcccggatc tgtgcttgcc caggaagaga caggaaggcg 780
 gatgaagata gcatcagaaa gcagcaagtt tcggacagta caaagaacgg tgatggtacg 840
 aagcgcctgt ttcgtcagaa cacacatggt atccagatga catccatcaa gaaacgaaga 900
 tcccagatg atgaactgtt ataactacca gtgaggggccc gtgagactta tgaaatgctg 960
 ttgaagatca aagagtcctt ggaactcatg cagtaccttc ctcagcacac aattgaaacg 1020
 tacaggcaac agcaacagca gcagcaccag cacttacttc agaaacagac ctcaatacag 1080
 tctccatctt catatggtaa cagctcccca cctctgaaca aaatgaacag catgaacaag 1140
 ctgccttctg tgagccagct tatcaacct cagcagcgca acgcctcac tcctacaacc 1200
 attcctgatg gcatgggagc caacattccc atgatgggca cccacatgcc aatggctgga 1260
 gacatgaatg gactcagccc caccagggca ctccctcccc cactctccat gccatccacc 1320
 tcccactgca cacccccacc tccgtatccc acagattgca gcattgtcag gatctggcaa 1380
 gtctga 1386

<210> 337
 <211> 1551
 <212> DNA
 <213> Homo sapiens

<400> 337
 atgtcccaga gcacacagac aaatgaattc ctcagtccag aggttttcca gcatatctgg 60
 gatthttctgg aacagcctat atgttcagtt cagcccattg acttgaactt tgtggatgaa 120
 ccatcagaag atggtgcgac aaacaagatt gagattagca tggactgtat ccgcatgcag 180
 gactcggacc tgagtgaacc catgtggcca cagtacacga acctggggct cctgaacagc 240
 atggaccagc agattcagaa cggtcctcg tccaccagtc cctataacac agaccacgcg 300
 cagaacagcg tcacggcgcc ctgcacctac gcacagccca gctccacctt cgatgctctc 360
 tctccatcac ccgccatccc ctccaacacc gactacccag gcccgcacag tttcgacgtg 420
 tcttccagc agtcgagcac cgccaagtgc gccacctgga cgtattccac tgaactgaag 480
 aaactctact gccaaattgc aaagacatgc cccatccaga tcaaggtgat gacccccact 540
 cctcagggag ctgttatccg cgccatgcct gtctacaaaa aagctgagca cgtcacggag 600
 gtggtgaagc ggtgccccaa ccatgagctg agccgtgaat tcaacgaggg acagattgcc 660
 cctcctagtc atttgattcg agtagagggg aacagccatg cccagtatgt agaagatccc 720
 atcacggaa gacagagtgt gctggtaacct tatgagccac cccaggttgg cactgaattc 780
 acgacagtct tgtacaattt catgtgtaac agcagttgtg ttggagggat gaaccgccgt 840

```

ccaatttttaa tcattgttac tctggaaacc agagatgggc aagtcctggg ccgacgctgc 900
tttgaggccc ggatctgtgc ttgccagga agagacagga aggcggatga agatagcatc 960
agaaagcagc aagtttcgga cagtacaaag aacgggtgatg gtacgaagcg cccgtttcgt 1020
cagaacacac atggtatcca gatgacatcc atcaagaaac gaagatcccc agatgatgaa 1080
ctgttatact taccagttag gggcctgag acttatgaaa tgctgttgaa gatcaaagag 1140
tccctggaac tcatgcagta ccttcctcag cacacaattg aaacgtacag gcaacagcaa 1200
cagcagcagc accagcactt acttcagaaa cagacctcaa tacagtctcc atcttcatat 1260
ggtaacagct cccacactct gaacaaaatg aacagcatga acaagctgcc ttctgtgagc 1320
cagcttatca accctcagca gcgcaacgcc ctcaactcta caaccattcc tgatggcatg 1380
ggagccaaca ttcccatgat gggcaccac atgccaatgg ctggagacat gaatggactc 1440
agccccaccc aggcactccc tccccactc tccatgccat ccacctccca ctgcacaccc 1500
ccacctccgt atccacaga ttgcagcatt gtcaggatct ggcaagtctg a 1551

```

<210> 338

<211> 586

<212> PRT

<213> Homo sapiens

<400> 338

```

Met Leu Tyr Leu Glu Asn Asn Ala Gln Thr Gln Phe Ser Glu Pro Gln
      5                      10                      15

Tyr Thr Asn Leu Gly Leu Leu Asn Ser Met Asp Gln Gln Ile Arg Asn
      20                      25                      30

Gly Ser Ser Ser Thr Ser Pro Tyr Asn Thr Asp His Ala Gln Asn Ser
      35                      40                      45

Val Thr Ala Pro Ser Pro Tyr Ala Gln Pro Ser Pro Thr Phe Asp Ala
      50                      55                      60

Leu Ser Pro Ser Pro Ala Ile Pro Ser Asn Thr Asp Tyr Pro Gly Pro
      65                      70                      75                      80

His Ser Ser Asp Val Ser Phe Gln Gln Ser Ser Thr Ala Lys Ser Ala
      85                      90                      95

Thr Trp Thr Tyr Ser Thr Glu Leu Lys Lys Leu Tyr Cys Gln Ile Ala
      100                     105                     110

Lys Thr Cys Pro Ile Gln Ile Lys Val Met Thr Pro Pro Pro Gln Gly
      115                     120                     125

Ala Val Ile Arg Ala Met Pro Val Tyr Lys Lys Ala Glu His Val Thr
      130                     135                     140

Glu Val Val Lys Arg Cys Pro Asn His Glu Leu Ser Arg Glu Phe Asn
      145                     150                     155                     160

Glu Gly Gln Ile Ala Pro Pro Ser His Leu Ile Arg Val Glu Gly Asn
      165                     170                     175

Ser His Ala Gln Tyr Val Glu Asp Pro Ile Thr Gly Arg Gln Ser Val
      180                     185                     190

```

Leu Val Pro Tyr Glu Pro Pro Gln Val Gly Thr Glu Phe Thr Thr Val
 195 200 205
 Leu Tyr Asn Phe Met Cys Asn Ser Ser Cys Val Gly Gly Met Asn Arg
 210 215 220
 Arg Pro Ile Leu Ile Ile Val Thr Leu Glu Thr Arg Asp Gly Gln Val
 225 230 235 240
 Leu Gly Arg Arg Cys Phe Glu Ala Arg Ile Cys Ala Cys Pro Gly Arg
 245 250 255
 Asp Arg Lys Ala Asp Glu Asp Ser Ile Arg Lys Gln Gln Val Ser Asp
 260 265 270
 Ser Thr Lys Asn Gly Asp Gly Thr Lys Arg Pro Phe Arg Gln Asn Thr
 275 280 285
 His Gly Ile Gln Met Thr Ser Ile Lys Lys Arg Arg Ser Pro Asp Asp
 290 295 300
 Glu Leu Leu Tyr Leu Pro Val Arg Gly Arg Glu Thr Tyr Glu Met Leu
 305 310 315 320
 Leu Lys Ile Lys Glu Ser Leu Glu Leu Met Gln Tyr Leu Pro Gln His
 325 330 335
 Thr Ile Glu Thr Tyr Arg Gln Gln Gln Gln Gln His Gln His Leu
 340 345 350
 Leu Gln Lys Gln Thr Ser Ile Gln Ser Pro Ser Ser Tyr Gly Asn Ser
 355 360 365
 Ser Pro Pro Leu Asn Lys Met Asn Ser Met Asn Lys Leu Pro Ser Val
 370 375 380
 Ser Gln Leu Ile Asn Pro Gln Gln Arg Asn Ala Leu Thr Pro Thr Thr
 385 390 395 400
 Ile Pro Asp Gly Met Gly Ala Asn Ile Pro Met Met Gly Thr His Met
 405 410 415
 Pro Met Ala Gly Asp Met Asn Gly Leu Ser Pro Thr Gln Ala Leu Pro
 420 425 430
 Pro Pro Leu Ser Met Pro Ser Thr Ser His Cys Thr Pro Pro Pro Pro
 435 440 445
 Tyr Pro Thr Asp Cys Ser Ile Val Ser Phe Leu Ala Arg Leu Gly Cys
 450 455 460
 Ser Ser Cys Leu Asp Tyr Phe Thr Thr Gln Gly Leu Thr Thr Ile Tyr
 465 470 475 480

Gln Ile Glu His Tyr Ser Met Asp Asp Leu Ala Ser Leu Lys Ile Pro
 485 490 495

Glu Gln Phe Arg His Ala Ile Trp Lys Gly Ile Leu Asp His Arg Gln
 500 505 510

Leu His Glu Phe Ser Ser Pro Ser His Leu Leu Arg Thr Pro Ser Ser
 515 520 525

Ala Ser Thr Val Ser Val Gly Ser Ser Glu Thr Arg Gly Glu Arg Val
 530 535 540

Ile Asp Ala Val Arg Phe Thr Leu Arg Gln Thr Ile Ser Phe Pro Pro
 545 550 555 560

Arg Asp Glu Trp Asn Asp Phe Asn Phe Asp Met Asp Ala Arg Arg Asn
 565 570 575

Lys Gln Gln Arg Ile Lys Glu Glu Gly Glu
 580 585

<210> 339

<211> 641

<212> PRT

<213> Homo sapiens

<400> 339

Met Ser Gln Ser Thr Gln Thr Asn Glu Phe Leu Ser Pro Glu Val Phe
 5 10 15

Gln His Ile Trp Asp Phe Leu Glu Gln Pro Ile Cys Ser Val Gln Pro
 20 25 30

Ile Asp Leu Asn Phe Val Asp Glu Pro Ser Glu Asp Gly Ala Thr Asn
 35 40 45

Lys Ile Glu Ile Ser Met Asp Cys Ile Arg Met Gln Asp Ser Asp Leu
 50 55 60

Ser Asp Pro Met Trp Pro Gln Tyr Thr Asn Leu Gly Leu Leu Asn Ser
 65 70 75 80

Met Asp Gln Gln Ile Gln Asn Gly Ser Ser Ser Thr Ser Pro Tyr Asn
 85 90 95

Thr Asp His Ala Gln Asn Ser Val Thr Ala Pro Ser Pro Tyr Ala Gln
 100 105 110

Pro Ser Ser Thr Phe Asp Ala Leu Ser Pro Ser Pro Ala Ile Pro Ser
 115 120 125

Asn Thr Asp Tyr Pro Gly Pro His Ser Phe Asp Val Ser Phe Gln Gln
 130 135 140

Ser Ser Thr Ala Lys Ser Ala Thr Trp Thr Tyr Ser Thr Glu Leu Lys
 145 150 155 160
 Lys Leu Tyr Cys Gln Ile Ala Lys Thr Cys Pro Ile Gln Ile Lys Val
 165 170 175
 Met Thr Pro Pro Pro Gln Gly Ala Val Ile Arg Ala Met Pro Val Tyr
 180 185 190
 Lys Lys Ala Glu His Val Thr Glu Val Val Lys Arg Cys Pro Asn His
 195 200 205
 Glu Leu Ser Arg Glu Phe Asn Glu Gly Gln Ile Ala Pro Pro Ser His
 210 215 220
 Leu Ile Arg Val Glu Gly Asn Ser His Ala Gln Tyr Val Glu Asp Pro
 225 230 235 240
 Ile Thr Gly Arg Gln Ser Val Leu Val Pro Tyr Glu Pro Pro Gln Val
 245 250 255
 Gly Thr Glu Phe Thr Thr Val Leu Tyr Asn Phe Met Cys Asn Ser Ser
 260 265 270
 Cys Val Gly Gly Met Asn Arg Arg Pro Ile Leu Ile Ile Val Thr Leu
 275 280 285
 Glu Thr Arg Asp Gly Gln Val Leu Gly Arg Arg Cys Phe Glu Ala Arg
 290 295 300
 Ile Cys Ala Cys Pro Gly Arg Asp Arg Lys Ala Asp Glu Asp Ser Ile
 305 310 315 320
 Arg Lys Gln Gln Val Ser Asp Ser Thr Lys Asn Gly Asp Gly Thr Lys
 325 330 335
 Arg Pro Phe Arg Gln Asn Thr His Gly Ile Gln Met Thr Ser Ile Lys
 340 345 350
 Lys Arg Arg Ser Pro Asp Asp Glu Leu Leu Tyr Leu Pro Val Arg Gly
 355 360 365
 Arg Glu Thr Tyr Glu Met Leu Leu Lys Ile Lys Glu Ser Leu Glu Leu
 370 375 380
 Met Gln Tyr Leu Pro Gln His Thr Ile Glu Thr Tyr Arg Gln Gln Gln
 385 390 395 400
 Gln Gln Gln His Gln His Leu Leu Gln Lys Gln Thr Ser Ile Gln Ser
 405 410 415
 Pro Ser Ser Tyr Gly Asn Ser Ser Pro Pro Leu Asn Lys Met Asn Ser
 420 425 430

Asn Ala Leu Thr Pro Thr Thr Ile Pro Asp Gly Met Gly Ala Asn Ile
450 455 460

Ser Pro Thr Gln Ala Leu Pro Pro Pro Leu Ser Met Pro Ser Thr Ser
485 490 495

His Cys Thr Pro Pro Pro Pro Tyr Pro Thr Asp Cys Ser Ile Val Gly
500 505 510

Phe Leu Ala Arg Leu Gly Cys Ser Ser Cys Leu Asp Tyr Phe Thr Thr
515 520 525

Gln Gly Leu Thr Thr Ile Tyr Gln Ile Glu His Tyr Ser Met Asp Asp
530 535 540

Leu Ala Ser Leu Lys Ile Pro Glu Gln Phe Arg His Ala Ile Trp Lys
545 550 555 560

Gly Ile Leu Asp His Arg Gln Leu His Glu Phe Ser Ser Pro Ser His
565 570 575

Leu Leu Arg Thr Pro Ser Ser Ala Ser Thr Val Ser Val Gly Ser Ser
580 585 590

Glu Thr Arg Gly Glu Arg Val Ile Asp Ala Val Arg Phe Thr Leu Arg
595 600 605

Gln Thr Ile Ser Phe Pro Pro Arg Asp Glu Trp Asn Asp Phe Asn Phe
610 615 620

Asp Met Asp Ala Arg Arg Asn Lys Gln Gln Arg Ile Lys Glu Glu Gly
625 630 635 640

Glu

<210> 340

<211> 448

<212> PRT

<213> Homo sapiens

<400> 340

Met Ser Gln Ser Thr Gln Thr Asn Glu Phe Leu Ser Pro Glu Val Phe
5 10 15

Gln His Ile Trp Asp Phe Leu Glu Gln Pro Ile Cys Ser Val Gln Pro
20 25 30

Ile Asp Leu Asn Phe Val Asp Glu Pro Ser Glu Asp Gly Ala Thr Asn
 35 40 45
 Lys Ile Glu Ile Ser Met Asp Cys Ile Arg Met Gln Asp Ser Asp Leu
 50 55 60
 Ser Asp Pro Met Trp Pro Gln Tyr Thr Asn Leu Gly Leu Leu Asn Ser
 65 70 75 80
 Met Asp Gln Gln Ile Gln Asn Gly Ser Ser Ser Thr Ser Pro Tyr Asn
 85 90 95
 Thr Asp His Ala Gln Asn Ser Val Thr Ala Pro Ser Pro Tyr Ala Gln
 100 105 110
 Pro Ser Ser Thr Phe Asp Ala Leu Ser Pro Ser Pro Ala Ile Pro Ser
 115 120 125
 Asn Thr Asp Tyr Pro Gly Pro His Ser Phe Asp Val Ser Phe Gln Gln
 130 135 140
 Ser Ser Thr Ala Lys Ser Ala Thr Trp Thr Tyr Ser Thr Glu Leu Lys
 145 150 155 160
 Lys Leu Tyr Cys Gln Ile Ala Lys Thr Cys Pro Ile Gln Ile Lys Val
 165 170 175
 Met Thr Pro Pro Pro Gln Gly Ala Val Ile Arg Ala Met Pro Val Tyr
 180 185 190
 Lys Lys Ala Glu His Val Thr Glu Val Val Lys Arg Cys Pro Asn His
 195 200 205
 Glu Leu Ser Arg Glu Phe Asn Glu Gly Gln Ile Ala Pro Pro Ser His
 210 215 220
 Leu Ile Arg Val Glu Gly Asn Ser His Ala Gln Tyr Val Glu Asp Pro
 225 230 235 240
 Ile Thr Gly Arg Gln Ser Val Leu Val Pro Tyr Glu Pro Pro Gln Val
 245 250 255
 Gly Thr Glu Phe Thr Thr Val Leu Tyr Asn Phe Met Cys Asn Ser Ser
 260 265 270
 Cys Val Gly Gly Met Asn Arg Arg Pro Ile Leu Ile Ile Val Thr Leu
 275 280 285
 Glu Thr Arg Asp Gly Gln Val Leu Gly Arg Arg Cys Phe Glu Ala Arg
 290 295 300
 Ile Cys Ala Cys Pro Gly Arg Asp Arg Lys Ala Asp Glu Asp Ser Ile
 305 310 315 320

Arg Lys Gln Gln Val Ser Asp Ser Thr Lys Asn Gly Asp Gly Thr Lys
325 330 335

Arg Pro Phe Arg Gln Asn Thr His Gly Ile Gln Met Thr Ser Ile Lys
340 345 350

Lys Arg Arg Ser Pro Asp Asp Glu Leu Leu Tyr Leu Pro Val Arg Gly
355 360 365

Arg Glu Thr Tyr Glu Met Leu Leu Lys Ile Lys Glu Ser Leu Glu Leu
370 375 380

Met Gln Tyr Leu Pro Gln His Thr Ile Glu Thr Tyr Arg Gln Gln Gln
385 390 395 400

Gln Gln Gln His Gln His Leu Leu Gln Lys His Leu Leu Ser Ala Cys
405 410 415

Phe Arg Asn Glu Leu Val Glu Pro Arg Arg Glu Thr Pro Lys Gln Ser
420 425 430

Asp Val Phe Phe Arg His Ser Lys Pro Pro Asn Arg Ser Val Tyr Pro
435 440 445

<210> 341

<211> 356

<212> PRT

<213> Homo sapiens

<400> 341

Met Leu Tyr Leu Glu Asn Asn Ala Gln Thr Gln Phe Ser Glu Pro Gln
5 10 15

Tyr Thr Asn Leu Gly Leu Leu Asn Ser Met Asp Gln Gln Ile Gln Asn
20 25 30

Gly Ser Ser Ser Thr Ser Pro Tyr Asn Thr Asp His Ala Gln Asn Ser
35 40 45

Val Thr Ala Pro Ser Pro Tyr Ala Gln Pro Ser Ser Thr Phe Asp Ala
50 55 60

Leu Ser Pro Ser Pro Ala Ile Pro Ser Asn Thr Asp Tyr Pro Gly Pro
65 70 75 80

His Ser Phe Asp Val Ser Phe Gln Gln Ser Ser Thr Ala Lys Ser Ala
85 90 95

Thr Trp Thr Tyr Ser Thr Glu Leu Lys Lys Leu Tyr Cys Gln Ile Ala
100 105 110

Lys Thr Cys Pro Ile Gln Ile Lys Val Met Thr Pro Pro Pro Gln Gly
115 120 125

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

Ala Val Ile Arg Ala Met Pro Val Tyr Lys Lys Ala Glu His Val Thr
 130 135 140

Glu Val Val Lys Arg Cys Pro Asn His Glu Leu Ser Arg Glu Phe Asn
 145 150 155 160

Glu Gly Gln Ile Ala Pro Pro Ser His Leu Ile Arg Val Glu Gly Asn
 165 170 175

Ser His Ala Gln Tyr Val Glu Asp Pro Ile Thr Gly Arg Gln Ser Val
 180 185 190

Leu Val Pro Tyr Glu Pro Pro Gln Val Gly Thr Glu Phe Thr Thr Val
 195 200 205

Leu Tyr Asn Phe Met Cys Asn Ser Ser Cys Val Gly Gly Met Asn Arg
 210 215 220

Arg Pro Ile Leu Ile Ile Val Thr Leu Glu Thr Arg Asp Gly Gln Val
 225 230 235 240

Leu Gly Arg Arg Cys Phe Glu Ala Arg Ile Cys Ala Cys Pro Gly Arg
 245 250 255

Asp Arg Lys Ala Asp Glu Asp Ser Ile Arg Lys Gln Gln Val Ser Asp
 260 265 270

Ser Thr Lys Asn Gly Asp Gly Thr Lys Arg Pro Ser Arg Gln Asn Thr
 275 280 285

His Gly Ile Gln Met Thr Ser Ile Lys Lys Arg Arg Ser Pro Asp Asp
 290 295 300

Glu Leu Leu Tyr Leu Pro Val Arg Gly Arg Glu Thr Tyr Glu Met Leu
 305 310 315 320

Leu Lys Ile Lys Glu Ser Leu Glu Leu Met Gln Tyr Leu Pro Gln His
 325 330 335

Thr Ile Glu Thr Tyr Arg Gln Gln Gln Gln Gln His Gln His Leu
 340 345 350

Leu Gln Lys Gln
 355

<210> 342

<211> 680

<212> PRT

<213> Homo sapiens

<400> 342

Met Asn Phe Glu Thr Ser Arg Cys Ala Thr Leu Gln Tyr Cys Pro Asp
 5 10 15

Pro	Tyr	Ile	Gln	Arg	Phe	Val	Glu	Thr	Pro	Ala	His	Phe	Ser	Trp	Lys	
			20						25			30				
Glu	Ser	Tyr	Tyr	Arg	Ser	Thr	Met	Ser	Gln	Ser	Thr	Gln	Thr	Asn	Glu	
			35			40						45				
Phe	Leu	Ser	Pro	Glu	Val	Phe	Gln	His	Ile	Trp	Asp	Phe	Leu	Glu	Gln	
50						55			60							
Pro	Ile	Cys	Ser	Val	Gln	Pro	Ile	Asp	Leu	Asn	Phe	Val	Asp	Glu	Pro	
65			70						75						80	
Ser	Glu	Asp	Gly	Ala	Thr	Asn	Lys	Ile	Glu	Ile	Ser	Met	Asp	Cys	Ile	
			85						90			95				
Arg	Met	Gln	Asp	Ser	Asp	Leu	Ser	Asp	Pro	Met	Trp	Pro	Gln	Tyr	Thr	
			100			105						110				
Asn	Leu	Gly	Leu	Leu	Asn	Ser	Met	Asp	Gln	Gln	Ile	Gln	Asn	Gly	Ser	
115						120						125				
Ser	Ser	Thr	Ser	Pro	Tyr	Asn	Thr	Asp	His	Ala	Gln	Asn	Ser	Val	Thr	
130						135			140							
Ala	Pro	Ser	Pro	Tyr	Ala	Gln	Pro	Ser	Ser	Thr	Phe	Asp	Ala	Leu	Ser	
145			150						155						160	
Pro	Ser	Pro	Ala	Ile	Pro	Ser	Asn	Thr	Asp	Tyr	Pro	Gly	Pro	His	Ser	
			165						170			175				
Phe	Asp	Val	Ser	Phe	Gln	Gln	Ser	Ser	Thr	Ala	Lys	Ser	Ala	Thr	Trp	
			180			185						190				
Thr	Tyr	Ser	Thr	Glu	Leu	Lys	Lys	Leu	Tyr	Cys	Gln	Ile	Ala	Lys	Thr	
195						200						205				
Cys	Pro	Ile	Gln	Ile	Lys	Val	Met	Thr	Pro	Pro	Pro	Gln	Gly	Ala	Val	
210						215			220							
Ile	Arg	Ala	Met	Pro	Val	Tyr	Lys	Lys	Ala	Glu	His	Val	Thr	Glu	Val	
225			230						235						240	
Val	Lys	Arg	Cys	Pro	Asn	His	Glu	Leu	Ser	Arg	Glu	Phe	Asn	Glu	Gly	
			245						250			255				
Gln	Ile	Ala	Pro	Pro	Ser	His	Leu	Ile	Arg	Val	Glu	Gly	Asn	Ser	His	
			260			265						270				
Ala	Gln	Tyr	Val	Glu	Asp	Pro	Ile	Thr	Gly	Arg	Gln	Ser	Val	Leu	Val	
275						280						285				
Pro	Tyr	Glu	Pro	Pro	Gln	Val	Gly	Thr	Glu	Phe	Thr	Thr	Val	Leu	Tyr	
290						295			300							

Asn Phe Met Cys Asn Ser Ser Cys Val Gly Gly Met Asn Arg Arg Pro
 305 310 315 320
 Ile Leu Ile Ile Val Thr Leu Glu Thr Arg Asp Gly Gln Val Leu Gly
 325 330 335
 Arg Arg Cys Phe Glu Ala Arg Ile Cys Ala Cys Pro Gly Arg Asp Arg
 340 345 350
 Lys Ala Asp Glu Asp Ser Ile Arg Lys Gln Gln Val Ser Asp Ser Thr
 355 360 365
 Lys Asn Gly Asp Gly Thr Lys Arg Pro Phe Arg Gln Asn Thr His Gly
 370 375 380
 Ile Gln Met Thr Ser Ile Lys Lys Arg Arg Ser Pro Asp Asp Glu Leu
 385 390 395 400
 Leu Tyr Leu Pro Val Arg Gly Arg Glu Thr Tyr Glu Met Leu Leu Lys
 405 410 415
 Ile Lys Glu Ser Leu Glu Leu Met Gln Tyr Leu Pro Gln His Thr Ile
 420 425 430
 Glu Thr Tyr Arg Gln Gln Gln Gln Gln His Gln His Leu Leu Gln
 435 440 445
 Lys Gln Thr Ser Ile Gln Ser Pro Ser Ser Tyr Gly Asn Ser Ser Pro
 450 455 460
 Pro Leu Asn Lys Met Asn Ser Met Asn Lys Leu Pro Ser Val Ser Gln
 465 470 475 480
 Leu Ile Asn Pro Gln Gln Arg Asn Ala Leu Thr Pro Thr Thr Ile Pro
 485 490 495
 Asp Gly Met Gly Ala Asn Ile Pro Met Met Gly Thr His Met Pro Met
 500 505 510
 Ala Gly Asp Met Asn Gly Leu Ser Pro Thr Gln Ala Leu Pro Pro Pro
 515 520 525
 Leu Ser Met Pro Ser Thr Ser Gln Cys Thr Pro Pro Pro Pro Tyr Pro
 530 535 540
 Thr Asp Cys Ser Ile Val Ser Phe Leu Ala Arg Leu Gly Cys Ser Ser
 545 550 555 560
 Cys Leu Asp Tyr Phe Thr Thr Gln Gly Leu Thr Thr Ile Tyr Gln Ile
 565 570 575
 Glu His Tyr Ser Met Asp Asp Leu Ala Ser Leu Lys Ile Pro Glu Gln
 580 585 590

Phe Arg His Ala Ile Trp Lys Gly Ile Leu Asp His Arg Gln Leu His
595 600 605

Glu Phe Ser Ser Pro Ser His Leu Leu Arg Thr Pro Ser Ser Ala Ser
610 615 620

Thr Val Ser Val Gly Ser Ser Glu Thr Arg Gly Glu Arg Val Ile Asp
625 630 635 640

Ala Val Arg Phe Thr Leu Arg Gln Thr Ile Ser Phe Pro Pro Arg Asp
645 650 655

Glu Trp Asn Asp Phe Asn Phe Asp Met Asp Ala Arg Arg Asn Lys Gln
660 665 670

Gln Arg Ile Lys Glu Glu Gly Glu
675 680

<210> 343

<211> 461

<212> PRT

<213> Homo sapiens

<400> 343

Met Leu Tyr Leu Glu Asn Asn Ala Gln Thr Gln Phe Ser Glu Pro Gln
5 10 15

Tyr Thr Asn Leu Gly Leu Leu Asn Ser Met Asp Gln Gln Ile Gln Asn
20 25 30

Gly Ser Ser Ser Thr Ser Pro Tyr Asn Thr Asp His Ala Gln Asn Ser
35 40 45

Val Thr Ala Pro Ser Pro Tyr Ala Gln Pro Ser Ser Thr Phe Asp Ala
50 55 60

Leu Ser Pro Ser Pro Ala Ile Pro Ser Asn Thr Asp Tyr Pro Gly Pro
65 70 75 80

His Ser Phe Asp Val Ser Phe Gln Gln Ser Ser Thr Ala Lys Ser Ala
85 90 95

Thr Trp Thr Tyr Ser Thr Glu Leu Lys Lys Leu Tyr Cys Gln Ile Ala
100 105 110

Lys Thr Cys Pro Ile Gln Ile Lys Val Met Thr Pro Pro Pro Gln Gly
115 120 125

Ala Val Ile Arg Ala Met Pro Val Tyr Lys Lys Ala Glu His Val Thr
130 135 140

Glu Val Val Lys Arg Cys Pro Asn His Glu Leu Ser Arg Glu Phe Asn
145 150 155 160

164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

Glu Gly Gln Ile Ala Pro Pro Ser His Leu Ile Arg Val Glu Gly Asn
 165 170 175
 Ser His Ala Gln Tyr Val Glu Asp Pro Ile Thr Gly Arg Gln Ser Val
 180 185 190
 Leu Val Pro Tyr Glu Pro Pro Gln Val Gly Thr Glu Phe Thr Thr Val
 195 200 205
 Leu Tyr Asn Phe Met Cys Asn Ser Ser Cys Val Gly Gly Met Asn Arg
 210 215 220
 Arg Pro Ile Leu Ile Ile Val Thr Leu Glu Thr Arg Asp Gly Gln Val
 225 230 235 240
 Leu Gly Arg Arg Cys Phe Glu Ala Arg Ile Cys Ala Cys Pro Gly Arg
 245 250 255
 Asp Arg Lys Ala Asp Glu Asp Ser Ile Arg Lys Gln Gln Val Ser Asp
 260 265 270
 Ser Thr Lys Asn Gly Asp Gly Thr Lys Arg Pro Phe Arg Gln Asn Thr
 275 280 285
 His Gly Ile Gln Met Thr Ser Ile Lys Lys Arg Arg Ser Pro Asp Asp
 290 295 300
 Glu Leu Leu Tyr Leu Pro Val Arg Gly Arg Glu Thr Tyr Glu Met Leu
 305 310 315 320
 Leu Lys Ile Lys Glu Ser Leu Glu Leu Met Gln Tyr Leu Pro Gln His
 325 330 335
 Thr Ile Glu Thr Tyr Arg Gln Gln Gln Gln Gln His Gln His Leu
 340 345 350
 Leu Gln Lys Gln Thr Ser Ile Gln Ser Pro Ser Ser Tyr Gly Asn Ser
 355 360 365
 Ser Pro Pro Leu Asn Lys Met Asn Ser Met Asn Lys Leu Pro Ser Val
 370 375 380
 Ser Gln Leu Ile Asn Pro Gln Gln Arg Asn Ala Leu Thr Pro Thr Thr
 385 390 395 400
 Ile Pro Asp Gly Met Gly Ala Asn Ile Pro Met Met Gly Thr His Met
 405 410 415
 Pro Met Ala Gly Asp Met Asn Gly Leu Ser Pro Thr Gln Ala Leu Pro
 420 425 430
 Pro Pro Leu Ser Met Pro Ser Thr Ser His Cys Thr Pro Pro Pro Pro
 435 440 445

```
<210> 344
<211> 516
<212> PRT
<213> Homo sapiens
```

Met Ser Gln Ser Thr Gln Thr Asn Glu Phe Leu Ser Pro Glu Val Phe
5 10 15

Gln His Ile Trp Asp Phe Leu Glu Gln Pro Ile Cys Ser Val Gln Pro
20 25 30

Ile Asp Leu Asn Phe Val Asp Glu Pro Ser Glu Asp Gly Ala Thr Asn
35 40 45

Lys Ile Glu Ile Ser Met Asp Cys Ile Arg Met Gln Asp Ser Asp Leu
50 55 60

Ser Asp Pro Met Trp Pro Gln Tyr Thr Asn Leu Gly Leu Leu Asn Ser
65 70 75 80

Met Asp Gln Gln Ile Gln Asn Gly Ser Ser Ser Thr Ser Pro Tyr Asn
85 90 95

Thr Asp His Ala Gln Asn Ser Val Thr Ala Pro Ser Pro Tyr Ala Gln
100 105 110

Pro Ser Ser Thr Phe Asp Ala Leu Ser Pro Ser Pro Ala Ile Pro Ser
115 120 125

Asn Thr Asp Tyr Pro Gly Pro His Ser Phe Asp Val Ser Phe Gln Gln
130 135 140

Ser Ser Thr Ala Lys Ser Ala Thr Trp Thr Tyr Ser Thr Glu Leu Lys
145 150 155 160

Lys Leu Tyr Cys Gln Ile Ala Lys Thr Cys Pro Ile Gln Ile Lys Val
165 170 175

Met Thr Pro Pro Pro Gln Gly Ala Val Ile Arg Ala Met Pro Val Tyr
180 185 190

Lys Lys Ala Glu His Val Thr Glu Val Val Lys Arg Cys Pro Asn His
195 200 205

Glu Leu Ser Arg Glu Phe Asn Glu Gly Gln Ile Ala Pro Pro Ser His
210 215 220

Leu Ile Arg Val Glu Gly Asn Ser His Ala Gln Tyr Val Glu Asp Pro
225 230 235 240

Ile Thr Gly Arg Gln Ser Val Leu Val Pro Tyr Glu Pro Pro Gln Val
245 250 255

Gly Thr Glu Phe Thr Thr Val Leu Tyr Asn Phe Met Cys Asn Ser Ser
260 265 270

Cys Val Gly Gly Met Asn Arg Arg Pro Ile Leu Ile Ile Val Thr Leu
275 280 285

Glu Thr Arg Asp Gly Gln Val Leu Gly Arg Arg Cys Phe Glu Ala Arg
290 295 300

Ile Cys Ala Cys Pro Gly Arg Asp Arg Lys Ala Asp Glu Asp Ser Ile
305 310 315 320

Arg Lys Gln Gln Val Ser Asp Ser Thr Lys Asn Gly Asp Gly Thr Lys
325 330 335

Arg Pro Phe Arg Gln Asn Thr His Gly Ile Gln Met Thr Ser Ile Lys
340 345 350

Lys Arg Arg Ser Pro Asp Asp Glu Leu Leu Tyr Leu Pro Val Arg Gly
355 360 365

Arg Glu Thr Tyr Glu Met Leu Leu Lys Ile Lys Glu Ser Leu Glu Leu
370 375 380

Met Gln Tyr Leu Pro Gln His Thr Ile Glu Thr Tyr Arg Gln Gln Gln
385 390 395 400

Gln Gln Gln His Gln His Leu Leu Gln Lys Gln Thr Ser Ile Gln Ser
405 410 415

Pro Ser Ser Tyr Gly Asn Ser Ser Pro Pro Leu Asn Lys Met Asn Ser
420 425 430

Met Asn Lys Leu Pro Ser Val Ser Gln Leu Ile Asn Pro Gln Gln Arg
435 440 445

Asn Ala Leu Thr Pro Thr Thr Ile Pro Asp Gly Met Gly Ala Asn Ile
450 455 460

Pro Met Met Gly Thr His Met Pro Met Ala Gly Asp Met Asn Gly Leu
465 470 475 480

Ser Pro Thr Gln Ala Leu Pro Pro Pro Leu Ser Met Pro Ser Thr Ser
485 490 495

His Cys Thr Pro Pro Pro Pro Tyr Pro Thr Asp Cys Ser Ile Val Arg
500 505 510

Ile Trp Gln Val
515

167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

<210> 345
 <211> 1800
 <212> DNA
 <213> Homo sapiens

<400> 345
 gcgcctcatt gccactgcag tgactaaagc tgggaagacg ctggtcagtt cacctgcccc 60
 actggttggt ttttaaaca attctgatac aggcgacatc ctactgacc gagcaaagat 120
 tgacattcgt atcatcactg tgcaccattg gcttctaggc actccagtgg ggtaggagaa 180
 ggaggtctga aaccctcgca gagggatctt gccctcattc tttgggtctg aaacactggc 240
 agtcgttgga aacaggactc agggataaac cagcgcaatg gattggggga cgctgcacac 300
 tttcatcggg ggtgtcaaca aacactccac cagcatcggg aaggtgtgga tcacagtcac 360
 ctttattttc cgagtcacga tcctagtggg ggctgccag gaagtgtggg gtgacgagca 420
 agaggacttc gtcgcaaca cactgcaacc gggatgcaaa aatgtgtgct atgaccactt 480
 tttcccggtg tcccacatcc ggctgtgggc cctccagctg atcttctgtc ccaccccagc 540
 gctgctggtg gccatgcacg tggcctaact caggcacgaa accactcgca agttcaggcg 600
 aggagagaag aggaatgatt tcaaagacat agaggacatt aaaaagcaca aggttcggat 660
 agaggggtcg ctgtgggtgga cgtacaccag cagcatcttt ttccgaatca tctttgaagc 720
 agcctttatg tatgtgtttt acttccttta caatgggtac cacctgccct ggggtgtgaa 780
 atgtgggatt gacccctgcc ccaaccttgt tgactgcttt atttctaggc caacagagaa 840
 gaccgtgttt accattttta tgatttctgc gtctgtgatt tgcattgctg ttaacgtggc 900
 agagttgtgc tacctgctgc tgaaagtgtg ttttaggaga tcaaagagag cacagacgca 960
 aaaaaatcac cccaatcatg ccctaaagga gagtaagcag aatgaaatga atgagctgat 1020
 ttcagatagt ggtcaaaatg caatcacagg tttcccaagc taaacatttc aaggtaaaat 1080
 gtagctgcgt cataaggaga cttctgtctt ctccagaagg caataccaac ctgaaagtgc 1140
 cttctgtagc ctgaagagtt tgtaaatgac tttcataata aatagacact tgagttaact 1200
 tttttagtaga tacttgctcc attcatacac aacgtaataa aatatgtggg ccatctctga 1260
 aaacaagaga ctgcttgaca aaggagcatt gcagtcactt tgacagggtc cttttaagtg 1320
 gactctctga caaagtgggt actttctgaa aatttatata actgttggtg ataaggaaca 1380
 tttatccagg aattgatacg tttattagga aaagatatat ttataggctt ggatgttttt 1440
 agttccgact ttgaatttat ataaagtatt tttataatga ctggtcttcc ttacctggaa 1500
 aaacatgcga tgttagtttt agaattacac cacaagtatc taaatttcca acttacaag 1560
 ggctctatct tgtaaatatt gttttgcatt gtctgttggc aaatttgtga actgtcatga 1620
 tacgcttaag gtgggaaagt gttcattgca caatatattt ttactgcttt ctgaatgtag 1680
 acggaacagt gtggaagcag aaggcttttt taactcatcc gtttgccga tcgttcgaga 1740
 ccaactgggag atgtggatgt ggttgccctc ttttgctcgt ccccggtggc taaccttct 1800

<210> 346
 <211> 261
 <212> PRT
 <213> Homo sapiens

<400> 346
 Met Asp Trp Gly Thr Leu His Thr Phe Ile Gly Gly Val Asn Lys His
 5 10 15
 Ser Thr Ser Ile Gly Lys Val Trp Ile Thr Val Ile Phe Ile Phe Arg
 20 25 30
 Val Met Ile Leu Val Val Ala Ala Gln Glu Val Trp Gly Asp Glu Gln

35	40	45
Glu Asp Phe Val Cys Asn Thr Leu Gln Pro Gly Cys Lys Asn Val Cys		
50	55	60
Tyr Asp His Phe Phe Pro Val Ser His Ile Arg Leu Trp Ala Leu Gln		
65	70	80
Leu Ile Phe Val Ser Thr Pro Ala Leu Leu Val Ala Met His Val Ala		
	85	90
Tyr Tyr Arg His Glu Thr Thr Arg Lys Phe Arg Arg Gly Glu Lys Arg		
	100	110
Asn Asp Phe Lys Asp Ile Glu Asp Ile Lys Lys His Lys Val Arg Ile		
	115	125
Glu Gly Ser Leu Trp Trp Thr Tyr Thr Ser Ser Ile Phe Phe Arg Ile		
	130	140
Ile Phe Glu Ala Ala Phe Met Tyr Val Phe Tyr Phe Leu Tyr Asn Gly		
	145	155
Tyr His Leu Pro Trp Val Leu Lys Cys Gly Ile Asp Pro Cys Pro Asn		
	165	175
Leu Val Asp Cys Phe Ile Ser Arg Pro Thr Glu Lys Thr Val Phe Thr		
	180	190
Ile Phe Met Ile Ser Ala Ser Val Ile Cys Met Leu Leu Asn Val Ala		
	195	205
Glu Leu Cys Tyr Leu Leu Leu Lys Val Cys Phe Arg Arg Ser Lys Arg		
	210	220
Ala Gln Thr Gln Lys Asn His Pro Asn His Ala Leu Lys Glu Ser Lys		
	225	235
Gln Asn Glu Met Asn Glu Leu Ile Ser Asp Ser Gly Gln Asn Ala Ile		
	245	255
Thr Gly Phe Pro Ser		
	260	

<210> 347

<211> 1740

<212> DNA

<213> Homo sapiens

<400> 347

atgaacaaac tgtatatcgg aaacctcagc gagaacgccg cccctcgga cctagaaagt 60
atcttcaagg acgccaagat cccggtgtcg ggacccttcc tgggtgaagac tgggtacgcg 120

```

ttcgtggact gcccggacga gagctgggcc ctcaaggcca tcgaggcgct ttcaggtaaa 180
atagaactgc acgggaaaacc catagaagtt gagcactcgg tcccaaaaag gcaaaggatt 240
cggaacttc agatacgaat tatcccgcct catttacagt gggaggtgct ggatagttta 300
ctagtccagt atggagtggt ggagagctgt gagcaagtga aactgactc ggaaactgca 360
gttgtaaagt taacctattc cagtaaggac caagctagac aagcactaga caaactgaat 420
ggatttcagt tagagaattt caccttgaaa gtagcctata tccctgatga aacggccgcc 480
cagcaaaacc ccttgacgca gcccggaggt cgcggggggc ttgggcagag gggctcctca 540
aggcaggggt ctccaggatc cgtatccaag cagaaacccat gtgatttgcc tctgcgcctg 600
ctggttccca cccaatttgt tggagccatc ataggaaaag aaggtgccac cattcggaac 660
atcaccaaac agaccagtc taaaatcgat gtccaccgta aagaaaatgc gggggctgct 720
gagaagtcga ttactatcct ctctactcct gaaggcacct ctgcggcttg taagtctatt 780
ctggagatta tgcataagga agctcaagat ataaaattca cagaagagat ccccttgaag 840
atatttagtc ataataactt tgttgagcgt cttattggta aagaaggaag aaatcttaa 900
aaaattgagc aagacacaga cactaaaatc acgatatctc cattgcagga attgacgctg 960
tataatccag aacgcactat tacagttaaa ggcaatgttg agacatgtgc caaagctgag 1020
gaggagatca tgaagaaaat cagggagctt tatgaaaatg atattgcttc tatgaatctt 1080
caagcacatt taattcctgg attaaatctg aacgccttgg gtctgttccc acccaattca 1140
gggatgccac ctcccacctc agggccccc tccagccatga ctccctccta cccgcagttt 1200
gagcaatcag aaacggagac tgttcactct tttatcccag ctctatcagt cggtgccatc 1260
atcggaagc aggggccagca catcaagcag ctttctcgtt ttgctggagc ttcaattaag 1320
attgctccag cggaagcacc agatgctaaa gtgaggatgg tgattatcac tggaccacca 1380
gaggctcagt tcaaggtcca gggaagaatt tatggaaaaa ttaaagaaga aaactttgtt 1440
agtcctaaag aagaggtgaa acttgaagct catatcagag tgccatcctt tgctgctggc 1500
agagttattg gaaaaggagg caaacgggtg aatgaacttc agaatttgtc aagtgcagaa 1560
gttggtgtcc ctgctgacca gacacctgat gagaatgacc aagtgtgtgt caaaataact 1620
ggtcacttct atgcttgcca ggttgcccag agaaaaattc aggaatttct gactcaggta 1680
aagcagcacc aacaacagaa ggctctgcaa agtgaccac ctcaagtcaag acggaagtaa 1740

```

<210> 348

<211> 579

<212> PRT

<213> Homo sapiens

<400> 348

```

Met Asn Lys Leu Tyr Ile Gly Asn Leu Ser Glu Asn Ala Ala Pro Ser
      5                      10                      15

```

```

Asp Leu Glu Ser Ile Phe Lys Asp Ala Lys Ile Pro Val Ser Gly Pro
      20                      25                      30

```

```

Phe Leu Val Lys Thr Gly Tyr Ala Phe Val Asp Cys Pro Asp Glu Ser
      35                      40                      45

```

```

Trp Ala Leu Lys Ala Ile Glu Ala Leu Ser Gly Lys Ile Glu Leu His
      50                      55                      60

```

```

Gly Lys Pro Ile Glu Val Glu His Ser Val Pro Lys Arg Gln Arg Ile
      65                      70                      75                      80

```

```

Arg Lys Leu Gln Ile Arg Asn Ile Pro Pro His Leu Gln Trp Glu Val
      85                      90                      95

```

```

Leu Asp Ser Leu Leu Val Gln Tyr Gly Val Val Glu Ser Cys Glu Gln

```

100	105	110
Val Asn Thr Asp Ser Glu Thr	Ala Val Val Asn Val Thr	Tyr Ser Ser
115	120	125
Lys Asp Gln Ala Arg Gln Ala	Leu Asp Lys Leu Asn Gly Phe	Gln Leu
130	135	140
Glu Asn Phe Thr Leu Lys Val	Ala Tyr Ile Pro Asp Glu Thr	Ala Ala
145	150	155 160
Gln Gln Asn Pro Leu Gln Gln	Pro Arg Gly Arg Arg Gly	Leu Gly Gln
165	170	175
Arg Gly Ser Ser Arg Gln Gly	Ser Pro Gly Ser Val Ser	Lys Gln Lys
180	185	190
Pro Cys Asp Leu Pro Leu Arg	Leu Leu Val Pro Thr Gln	Phe Val Gly
195	200	205
Ala Ile Ile Gly Lys Glu Gly	Ala Thr Ile Arg Asn Ile	Thr Lys Gln
210	215	220
Thr Gln Ser Lys Ile Asp Val	His Arg Lys Glu Asn Ala	Gly Ala Ala
225	230	235 240
Glu Lys Ser Ile Thr Ile Leu	Ser Thr Pro Glu Gly Thr	Ser Ala Ala
245	250	255
Cys Lys Ser Ile Leu Glu Ile	Met His Lys Glu Ala Gln	Asp Ile Lys
260	265	270
Phe Thr Glu Glu Ile Pro Leu	Lys Ile Leu Ala His Asn	Asn Phe Val
275	280	285
Gly Arg Leu Ile Gly Lys Glu	Gly Arg Asn Leu Lys Lys	Ile Glu Gln
290	295	300
Asp Thr Asp Thr Lys Ile Thr	Ile Ser Pro Leu Gln Glu	Leu Thr Leu
305	310	315 320
Tyr Asn Pro Glu Arg Thr Ile	Thr Val Lys Gly Asn Val	Glu Thr Cys
325	330	335
Ala Lys Ala Glu Glu Glu Ile	Met Lys Lys Ile Arg Glu	Ser Tyr Glu
340	345	350
Asn Asp Ile Ala Ser Met Asn	Leu Gln Ala His Leu Ile	Pro Gly Leu
355	360	365
Asn Leu Asn Ala Leu Gly Leu	Phe Pro Pro Thr Ser Gly	Met Pro Pro
370	375	380
Pro Thr Ser Gly Pro Pro Ser	Ala Met Thr Pro Pro Tyr	Pro Gln Phe

```

385              390              395              400
Glu Gln Ser Glu Thr Glu Thr Val His Leu Phe Ile Pro Ala Leu Ser
              405              410              415
Val Gly Ala Ile Ile Gly Lys Gln Gly Gln His Ile Lys Gln Leu Ser
              420              425              430
Arg Phe Ala Gly Ala Ser Ile Lys Ile Ala Pro Ala Glu Ala Pro Asp
              435              440              445
Ala Lys Val Arg Met Val Ile Ile Thr Gly Pro Pro Glu Ala Gln Phe
              450              455              460
Lys Ala Gln Gly Arg Ile Tyr Gly Lys Ile Lys Glu Glu Asn Phe Val
465              470              475              480
Ser Pro Lys Glu Glu Val Lys Leu Glu Ala His Ile Arg Val Pro Ser
              485              490              495
Phe Ala Ala Gly Arg Val Ile Gly Lys Gly Gly Lys Thr Val Asn Glu
              500              505              510
Leu Gln Asn Leu Ser Ser Ala Glu Val Val Val Pro Arg Asp Gln Thr
              515              520              525
Pro Asp Glu Asn Asp Gln Val Val Val Lys Ile Thr Gly His Phe Tyr
              530              535              540
Ala Cys Gln Val Ala Gln Arg Lys Ile Gln Glu Ile Leu Thr Gln Val
545              550              555              560
Lys Gln His Gln Gln Gln Lys Ala Leu Gln Ser Gly Pro Pro Gln Ser
              565              570              575
Arg Arg Lys

```

<210> 349

<211> 207

<212> DNA

<213> Homo sapiens

<400> 349

```

atgtggcagc cctctttctt caagtggctc ttgtcctggt gccctgggag ttctcaaatt 60
gctgcagcag cctccaccca gcctgaggat gacatcaata cacagaggaa gaagagtcag 120
gaaaagatga gagaagttac agactctcct gggcgacccc gagagcttac cattcctcag 180
acttcttcac atggtgctaa cagattt

```

207

<210> 350

<211> 69

<212> PRT

<213> Homo sapiens

<400> 350

Met Trp Gln Pro Leu Phe Phe Lys Trp Leu Leu Ser Cys Cys Pro Gly
 5 10 15

Ser Ser Gln Ile Ala Ala Ala Ala Ser Thr Gln Pro Glu Asp Asp Ile
 20 25 30

Asn Thr Gln Arg Lys Lys Ser Gln Glu Lys Met Arg Glu Val Thr Asp
 35 40 45

Ser Pro Gly Arg Pro Arg Glu Leu Thr Ile Pro Gln Thr Ser Ser His
 50 55 60

Gly Ala Asn Arg Phe
 65

<210> 351

<211> 1012

<212> DNA

<213> Homo sapiens

<400> 351

```

ccctctagaa ataattttgt ttaactttta gaaggagata tacatatgca tcaccatcac 60
catcacacgg ccgcgtccga taacttccag ctgtcccagg gtgggcaggg attcgccatt 120
ccgatcgggc aggcgatggc gatcgcgggc cagatcaagc ttcccaccgt tcatatcggg 180
cctaccgcct tctcgggctt ggggtgttgc gacaacaacg gcaacggcgc acgagtccaa 240
cgcggtggtc ggagcgctcc ggcggaagtc ctcggcattc ccaccggcga cgtgatcacc 300
gcggtcgacg gcgctccgat caactcggcc accgcgatgg cggacgcgct taacgggcat 360
catcccgtg acgtcatctc ggtgacctgg caaaccaagt cgggcggcac gcgtacaggg 420
aacgtgacat tggccgaggg acccccggcc gaattcatgg attgggggac gctgcacact 480
ttcatcgggg gtgtcaacaa aactcaccac agcatcggga aggtgtggat cacagtcata 540
tttattttcc gagtcatgat cctcgtggtg gctgccagg aagtgtgggg tgacgagcaa 600
gaggacttcg tctgcaaac actgcaaccg ggatgcaaaa atgtgtgcta tgaccacttt 660
ttcccgtgtg ccacatccg gctgtgggct ctcagctga tcttcgtctc caccocagcg 720
ctgctggtgg ccatgcatgt ggcctactac aggcacgaaa ccactcgcaa gttcaggcga 780
ggagagaaga ggaatgattt caaagacata gaggacatta aaaagcagaa gggtcggata 840
gaggggtgac tcgagcacca ccaccaccac cactgagatc cggctgctaa caaagccga 900
aagggaagctg agttggctgc tgccaccgct gagcaataac tagcataacc ccttggggcc 960
tctaaacggg tcttgagggg ttttttgcgt aaaggaggaa ctatatccgg at 1012

```

<210> 352

<211> 267

<212> PRT

<213> Homo sapiens

<400> 352

Met His His His His His Thr Ala Ala Ser Asp Asn Phe Gln Leu
 5 10 15

Ser Gln Gly Gly Gln Gly Phe Ala Ile Pro Ile Gly Gln Ala Met Ala
 20 25 30

Ile Ala Gly Gln Ile Lys Leu Pro Thr Val His Ile Gly Pro Thr Ala
 35 40 45

Phe Leu Gly Leu Gly Val Val Asp Asn Asn Gly Asn Gly Ala Arg Val
 50 55 60

Gln Arg Val Val Gly Ser Ala Pro Ala Ala Ser Leu Gly Ile Ser Thr
 65 70 75 80

Gly Asp Val Ile Thr Ala Val Asp Gly Ala Pro Ile Asn Ser Ala Thr
 85 90 95

Ala Met Ala Asp Ala Leu Asn Gly His His Pro Gly Asp Val Ile Ser
 100 105 110

Val Thr Trp Gln Thr Lys Ser Gly Gly Thr Arg Thr Gly Asn Val Thr
 115 120 125

Leu Ala Glu Gly Pro Pro Ala Glu Phe Met Asp Trp Gly Thr Leu His
 130 135 140

Thr Phe Ile Gly Gly Val Asn Lys His Ser Thr Ser Ile Gly Lys Val
 145 150 155 160

Trp Ile Thr Val Ile Phe Ile Phe Arg Val Met Ile Leu Val Val Ala
 165 170 175

Ala Gln Glu Val Trp Gly Asp Glu Gln Glu Asp Phe Val Cys Asn Thr
 180 185 190

Leu Gln Pro Gly Cys Lys Asn Val Cys Tyr Asp His Phe Phe Pro Val
 195 200 205

Ser His Ile Arg Leu Trp Ala Leu Gln Leu Ile Phe Val Ser Thr Pro
 210 215 220

Ala Leu Leu Val Ala Met His Val Ala Tyr Tyr Arg His Glu Thr Thr
 225 230 235 240

Arg Lys Phe Arg Arg Gly Glu Lys Arg Asn Asp Phe Lys Asp Ile Glu
 245 250 255

Asp Ile Lys Lys Gln Lys Val Arg Ile Glu Gly
 260 265

<210> 353

<211> 900

<212> DNA

<213> Homo sapiens

<400> 353

atgcatcacc atcaccatca cacggccgcg tccgataact tccagctgtc ccagggtggg 60


```

cagggattcg ccattccgat cgggcaggcg atggcgatcg cgggccagat caagcttccc 120
accgttcata tcgggcctac cgccttcctc ggcttggtg ttgtcgacaa caacggcaac 180
ggcgacgagag tccaacgcgt ggtcgggagc gctcggcg caagtctcg catctccacc 240
ggcgacgtga tcaccgcggt cgacggcgct ccgatcaact cggccaccgc gatggcgag 300
gcgcttaacg ggcatcatcc cggtgacgtc atctcggtga cctggcaaac caagtcgggc 360
ggcacgcgta cagggaaagt gacattggcc gagggacccc cggccgaatt ccacgaaacc 420
actcgcaagt tcaggcgagg agagaagagg aatgatttca aagacataga ggacattaaa 480
aagcagaagg ttcggataga ggggtcgctg tgggtggacgt acaccagcag catctttttc 540
cgaatcatct ttgaagcagc ctttatgtat gtgttttact tcctttacaa tgggtaccac 600
ctgccctggg tgttgaaatg tgggattgac cctgccccca accttggtga ctgctttatt 660
tctaggccaa cagagaagac cgtggttacc atttttatga tttctgcgtc tgtgatttgc 720
atgctgctta acgtggcaga gttgtgctac ctgctgctga aagtgtgttt taggagatca 780
aagagagcac agacgcaaaa aaatcacccc aatcatgccc taaaggagag taagcagaat 840
gaaatgaatg agctgatttc agatagtggg caaaatgcaa tcacagggtt cccaagctaa 900

```

<210> 354

<211> 299

<212> PRT

<213> Homo sapiens

<400> 354

```

Met His His His His His His Thr Ala Ala Ser Asp Asn Phe Gln Leu
      5                      10                      15

Ser Gln Gly Gly Gln Gly Phe Ala Ile Pro Ile Gly Gln Ala Met Ala
      20                      25                      30

Ile Ala Gly Gln Ile Lys Leu Pro Thr Val His Ile Gly Pro Thr Ala
      35                      40                      45

Phe Leu Gly Leu Gly Val Val Asp Asn Asn Gly Asn Gly Ala Arg Val
      50                      55                      60

Gln Arg Val Val Gly Ser Ala Pro Ala Ala Ser Leu Gly Ile Ser Thr
      65                      70                      75                      80

Gly Asp Val Ile Thr Ala Val Asp Gly Ala Pro Ile Asn Ser Ala Thr
      85                      90                      95

Ala Met Ala Asp Ala Leu Asn Gly His His Pro Gly Asp Val Ile Ser
      100                     105                     110

Val Thr Trp Gln Thr Lys Ser Gly Gly Thr Arg Thr Gly Asn Val Thr
      115                     120                     125

Leu Ala Glu Gly Pro Pro Ala Glu Phe His Glu Thr Thr Arg Lys Phe
      130                     135                     140

Arg Arg Gly Glu Lys Arg Asn Asp Phe Lys Asp Ile Glu Asp Ile Lys
      145                     150                     155                     160

Lys Gln Lys Val Arg Ile Glu Gly Ser Leu Trp Trp Thr Tyr Thr Ser
      165                     170                     175

```

Ser Ile Phe Phe Arg Ile Ile Phe Glu Ala Ala Phe Met Tyr Val Phe
 180 185 190

Tyr Phe Leu Tyr Asn Gly Tyr His Leu Pro Trp Val Leu Lys Cys Gly
 195 200 205

Ile Asp Pro Cys Pro Asn Leu Val Asp Cys Phe Ile Ser Arg Pro Thr
 210 215 220

Glu Lys Thr Val Phe Thr Ile Phe Met Ile Ser Ala Ser Val Ile Cys
 225 230 235 240

Met Leu Leu Asn Val Ala Glu Leu Cys Tyr Leu Leu Leu Lys Val Cys
 245 250 255

Phe Arg Arg Ser Lys Arg Ala Gln Thr Gln Lys Asn His Pro Asn His
 260 265 270

Ala Leu Lys Glu Ser Lys Gln Asn Glu Met Asn Glu Leu Ile Ser Asp
 275 280 285

Ser Gly Gln Asn Ala Ile Thr Gly Phe Pro Ser
 290 295

<210> 355

<211> 24

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer

<400> 355

ggagtacagc ttcaagacaa tggg

24

<210> 356

<211> 31

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer

<400> 356

ccatgggaat tcattataat aattttgttc c

31

<210> 357

<211> 920

<212> PRT

<213> Homo sapiens

<400> 357

Met Gln His His His His His His Gly Val Gln Leu Gln Asp Asn Gly

1	5	10	15
Tyr Asn Gly Leu Leu Ile Ala Ile Asn Pro Gln Val Pro Glu Asn Gln			
20	25	30	
Asn Leu Ile Ser Asn Ile Lys Glu Met Ile Thr Glu Ala Ser Phe Tyr			
35	40	45	
Leu Phe Asn Ala Thr Lys Arg Arg Val Phe Phe Arg Asn Ile Lys Ile			
50	55	60	
Leu Ile Pro Ala Thr Trp Lys Ala Asn Asn Asn Ser Lys Ile Lys Gln			
65	70	75	80
Glu Ser Tyr Glu Lys Ala Asn Val Ile Val Thr Asp Trp Tyr Gly Ala			
85	90	95	
His Gly Asp Asp Pro Tyr Thr Leu Gln Tyr Arg Gly Cys Gly Lys Glu			
100	105	110	
Gly Lys Tyr Ile His Phe Thr Pro Asn Phe Leu Leu Asn Asp Asn Leu			
115	120	125	
Thr Ala Gly Tyr Gly Ser Arg Gly Arg Val Phe Val His Glu Trp Ala			
130	135	140	
His Leu Arg Trp Gly Val Phe Asp Glu Tyr Asn Asn Asp Lys Pro Phe			
145	150	155	160
Tyr Ile Asn Gly Gln Asn Gln Ile Lys Val Thr Arg Cys Ser Ser Asp			
165	170	175	
Ile Thr Gly Ile Phe Val Cys Glu Lys Gly Pro Cys Pro Gln Glu Asn			
180	185	190	
Cys Ile Ile Ser Lys Leu Phe Lys Glu Gly Cys Thr Phe Ile Tyr Asn			
195	200	205	
Ser Thr Gln Asn Ala Thr Ala Ser Ile Met Phe Met Gln Ser Leu Ser			
210	215	220	
Ser Val Val Glu Phe Cys Asn Ala Ser Thr His Asn Gln Glu Ala Pro			
225	230	235	240
Asn Leu Gln Asn Gln Met Cys Ser Leu Arg Ser Ala Trp Asp Val Ile			
245	250	255	
Thr Asp Ser Ala Asp Phe His His Ser Phe Pro Met Asn Gly Thr Glu			
260	265	270	
Leu Pro Pro Pro Thr Phe Ser Leu Val Glu Ala Gly Asp Lys Val			
275	280	285	
Val Cys Leu Val Leu Asp Val Ser Ser Lys Met Ala Glu Ala Asp Arg			
290	295	300	
Leu Leu Gln Leu Gln Gln Ala Ala Glu Phe Tyr Leu Met Gln Ile Val			
305	310	315	320
Glu Ile His Thr Phe Val Gly Ile Ala Ser Phe Asp Ser Lys Gly Glu			
325	330	335	
Ile Arg Ala Gln Leu His Gln Ile Asn Ser Asn Asp Asp Arg Lys Leu			
340	345	350	
Leu Val Ser Tyr Leu Pro Thr Thr Val Ser Ala Lys Thr Asp Ile Ser			
355	360	365	
Ile Cys Ser Gly Leu Lys Lys Gly Phe Glu Val Val Glu Lys Leu Asn			
370	375	380	
Gly Lys Ala Tyr Gly Ser Val Met Ile Leu Val Thr Ser Gly Asp Asp			
385	390	395	400
Lys Leu Leu Gly Asn Cys Leu Pro Thr Val Leu Ser Ser Gly Ser Thr			
405	410	415	
Ile His Ser Ile Ala Leu Gly Ser Ser Ala Ala Pro Asn Leu Glu Glu			
420	425	430	
Leu Ser Arg Leu Thr Gly Gly Leu Lys Phe Phe Val Pro Asp Ile Ser			

865 870 875 880
 Lys Gly Val Leu Thr Ala Met Gly Leu Ile Gly Ile Ile Cys Leu Ile
 885 890 895
 Ile Val Val Thr His His Thr Leu Ser Arg Lys Lys Arg Ala Asp Lys
 900 905 910
 Lys Glu Asn Gly Thr Lys Leu Leu
 915 920

<210> 358
 <211> 2773
 <212> DNA
 <213> Homo sapiens

<400> 358
 catatgcagc atcaccacca tcaccacgga gtacagcttc aagacaatgg gtataatgga 60
 ttgctcattg caattaatcc tcaggtaacct gagaatcaga acctcatctc aaacattaag 120
 gaaatgataa ctgaagcttc attttaccta tttaatgcta ccaagagaag agtatttttc 180
 agaaatataa agatttttaac acctgccaca tggaaagcta ataataacag caaaataaaa 240
 caagaatcat atgaaaaggc aaatgtcata gtgactgact ggtatggggc acatggagat 300
 gatccataca ccctacaata cagaggggtgt ggaaaagagg gaaaatacat tcatttcaca 360
 cctaattttcc tactgaatga taacttaaca gctggctacg gatcacgagg ccgagtgttt 420
 gtccatgaat gggcccacct ccgttgggggt gtgttcgatg agtataacaa tgacaaacct 480
 ttctacataa atggggcaaaa tcaaattaaa gtgacaagggt gttcatctga catcacaggc 540
 attttttgtgt gtgaaaaagg tccttgcccc caagaaaact gtattattag taagcttttt 600
 aaagaaggat gcacctttat ctacaatagc acccaaatg caactgcac aataatgttc 660
 atgcaaagtt tatcttctgt ggttgaattt tgtaatgcaa gtaccacaa ccaagaagca 720
 ccaaacctac agaaccagat gtgcagcctc agaagtgcac gggatgtaat cacagactct 780
 gctgactttc accacagctt tcccatgaac gggactgagc ttccacctcc tccacattc 840
 tcgctttagt aggctgtgga caaagtggtc tgtttagtgc tggatgtgtc cagcaagatg 900
 gcagaggctg acagactcct tcaactacaa caagccgcag aattttattt gatgcagatt 960
 gttgaaattc ataccttcgt gggcattgcc agtttcgaca gcaaaggaga gatcagagcc 1020
 cagctacacc aaattaacag caatgatgat cgaaagttgc tggtttcata tctgcccacc 1080
 actgtatcag ctaaaacaga catcagcatt tgttcagggc ttaagaaagg atttgagggtg 1140
 gttgaaaaac tgaatggaaa agcttatggc tctgtgatga tattagtga cagcggagat 1200
 gataagcttc ttggcaattg cttaccacat gtgctcagca gtggttcaac aattcactcc 1260
 attgccctgg gtccatctgc agcccaaat ctggaggaat tatcacgtct tacaggaggt 1320
 ttaaagttct ttgttccaga tatatcaaac tccaatgaca tgattgatgc tttcagtaga 1380
 atttcctctg gaactggaga catttttocag caacatattc agcttgaaaag tacaggtgaa 1440
 aatgtcaaac ctcaccatca attgaaaaac acagtgactg tggataatac tgtgggcaac 1500
 gacactatgt ttctagttac gtggcaggcc agtggctctc ctgagattat attatttgat 1560
 cctgatggac gaaaatacta cacaaataat tttatcacca atctaacttt tcggacagct 1620
 agtctttgga ttccaggaac agctaagcct gggcactgga cttacacctt gaacaatacc 1680
 catcattctc tgcaagccct gaaagtgaca gtgacctctc gcgcctccaa ctcagctgtg 1740
 cccccagcca ctgtggaagc ctttgtggaa agagacagcc tccattttcc tcatcctgtg 1800
 atgatttatg ccaatgtgaa acagggattt tatcccatc ttaatgccac tgtcactgcc 1860
 acagttgagc cagagactgg agatcctgtt acgctgagac tccttgatga tggagcaggt 1920
 gctgatgtta taaaaaatga tggaaatttac tgcaggtatt ttttctcctt tgcagcaaat 1980
 ggtagatata gcttgaaagt gcatgtcaat cactctccca gcataagcac ccagcccac 2040
 tctattccag ggagtcatgc tatgtatgta ccaggttaca cagcaaacgg taatattcag 2100
 atgaatgtc caaggaaatc agtaggcaga aatgaggagg agcgaaagtg gggctttagc 2160
 cgagtcagct caggaggctc cttttcagtg ctgggagttc cagctggccc ccacctgat 2220
 gtgtttccac catgcaaaat tattgacctg gaagctgtaa aagtagaaga ggaattgacc 2280
 ctatcttgga cagcacctgg agaagacttt gatcagggcc aggtacaag ctatgaaata 2340
 agaatgagta aaagtctaca gaatatccaa gatgacttta acaatgctat tttagtaaat 2400

```

acatcaaagc gaaatcctca gcaagctggc atcagggaga tattttacgtt ctcaccccaa 2460
atttccacga atggacctga acatcagcca aatggagaaa cacatgaaag ccacagaatt 2520
tatgttgcaa tacgagcaat ggataggaac tccttacagt ctgctgtatc taacattgcc 2580
caggcgcttc tgtttattcc cccaattct gatcctgtac ctgccagaga ttatcttata 2640
ttgaaaggag ttttaacagc aatgggtttg ataggaatca tttgccttat tatagttgtg 2700
acacatcata cttaagcag gaaaaagaga gcagacaaga aagagaatgg aacaaaatta 2760
ttataatgaa ttc 2773

```

<210> 359

<211> 25

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer

<400> 359

tggcagcccc tcttcttcaa gtggc

25

<210> 360

<211> 33

<212> DNA

<213> Artificial Sequence

<220>

<223> Primer

<400> 360

cgccagaatt catcaaaca atctgttagc acc

33

<210> 361

<211> 77

<212> PRT

<213> Homo sapiens

<400> 361

```

Met Gln His His His His His His Trp Gln Pro Leu Phe Phe Lys Trp
  1              5              10              15
Leu Leu Ser Cys Cys Pro Gly Ser Ser Gln Ile Ala Ala Ala Ala Ser
      20              25              30
Thr Gln Pro Glu Asp Asp Ile Asn Thr Gln Arg Lys Lys Ser Gln Glu
      35              40              45
Lys Met Arg Glu Val Thr Asp Ser Pro Gly Arg Pro Arg Glu Leu Thr
      50              55              60
Ile Pro Gln Thr Ser Ser His Gly Ala Asn Arg Phe Val
65              70              75

```

<210> 362

<211> 244

<212> DNA

<213> Homo sapiens

<400> 362

catatgcagc atcaccacca tcaccactgg cagcccctct tcttcaagtg gctcttgctc

60

```

tgttgcoctg ggagttctca aattgctgca gcagcctcca cccagcctga ggatgacatc 120
aatacacaga ggaagaagag tcaggaaaag atgagagaag ttacagactc tcctgggcga 180
ccccgagagc ttaccattcc tcagacttct tcacatgggtg ctaacagatt tgtttgatga 240
attc 244

```

```

<210> 363
<211> 20
<212> PRT
<213> Homo sapiens

```

```

<400> 363
Met Trp Gln Pro Leu Phe Phe Lys Trp Leu Leu Ser Cys Cys Pro Gly
          5              10              15
Ser Ser Gln Ile
          20

```

```

<210> 364
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<400> 364
atgtggcagc ccctcttctt caagtggctc ttgtcctggt gccctgggag ttctcaaatt 60

```

```

<210> 365
<211> 20
<212> PRT
<213> Homo sapiens

```

```

<400> 365
Gly Ser Ser Gln Ile Ala Ala Ala Ala Ser Thr Gln Pro Glu Asp Asp
          5              10              15
Ile Asn Thr Gln
          20

```

```

<210> 366
<211> 60
<212> DNA
<213> Homo sapiens

```

```

<400> 366
gggagttctc aaattgctgc agcagcctcc acccagcctg aggatgacat caatacacag 60

```

```

<210> 367
<211> 20
<212> PRT
<213> Homo sapiens

```

```

<400> 367
Lys Pro Gly His Trp Thr Tyr Thr Leu Asn Asn Thr His His Ser Leu
          5              10              15
Gln Ala Leu Lys

```

20

<210> 368
 <211> 2343
 <212> DNA
 <213> Homo sapiens

<400> 368

```

attccggagc gtttgcggtc tcgcttcatg gccgctctcc cgccctcct gggatctgtg      60
gggagctggg gagcccgagc cggcccgag cggagctgg cgagccgagc ggagacctgt      120
gcgccgcgcc tctgaggcgc agcatgtgaa gcggagacgg catccagtgg ggggcgagcc      180
tctcagccgg ccgggatggc taccacggcc gagctcttcg aggagccttt tgtggcagat      240
gaatatattg aacgtcttgt atggagaacc ccaggaggag gctctagagg tggacctgaa      300
gcttttgatc ctaaaagatt attagaagaa tttgtaaadc atattcagga actccagata      360
atggatgaaa ggattcagag gaaagtagag aaactagagc aacaatgtca gaaagaagcc      420
aaggaatttg ccaagaaggt acaagagctg cagaaaagca atcagggttg cttccaacat      480
ttccaagaac tagatgagca cattagctat gtagcaacta aagtctgtca ccttgagagc      540
cagtttagag ggttaaaccac accagacaa cgggcagtg aggctcagaa attgatgaaa      600
tactttaatg agtttctaga tggagaattg aaatctgatg tttttacaaa ttctgaaaag      660
ataaaggaag cagcagacat cattcagaag ttgcacctaa ttgccaaga gttacctttt      720
gatagatttt cagaagttaa atccaaaatt gcaagtaaat accatgattt agaatgccag      780
ctgattcagg agtttaccag tgctcaaaga agaggtgaaa tctccagaat gagagaagta      840
gcagcagttt tacttcattt taagggttat tccattgtg ttgatgttta tataaagcag      900
tgccaggagg gtgcttattt gagaaatgat atatttgaag acgctggaat actctgtcaa      960
agagtgaaca aacaagttgg agatatcttc agtaatccag aaacagtcct ggctaaactt     1020
attcaaaatg tatttgaaat caaactacag agttttgtga aagagcagtt agaagaatgt     1080
aggaagtccg atgcagagca atatctcaaa aatctctatg atctgtatac aagaaccacc     1140
aatctttcca gcaagctgat ggagtttaat ttaggtactg ataaacagac tttcttgtct     1200
aagcttatca aatccatttt catttcctat ttggagaact atattgaggt ggagactgga     1260
tatttgaaaa gcagaagtgc tatgatccta cagcgtat atgattcgaa aaaccatcaa     1320
aagagatcca ttggcacagg aggtattcaa gatttgaagg aaagaattag acagcgtaac     1380
aacttaccac ttgggccaag tatcgatact catggggaga cttttctatc ccaagaagtg     1440
gtggttaatc ttttacaaga aaccaaaaca gcctttgaaa gatgtcatag gctctctgat     1500
ccttctgact taccaaaggaa tgccctcaga atttttacca ttcttgtgga atttttatgt     1560
attgagcata ttgattatgc tttggaaaca ggacttgctg gaattccctc ttcagattct     1620
aggaatgcaa atcttttatt tttggacgtt gtgcaacagg ccaatactat ttttcatctt     1680
tttgacaaac agtttaattga tcaccttatg ccactaataa gctcttctcc taagttatct     1740
gaatgccttc agaagaaaaa agaaataatt gaacaaatgg agatgaaatt ggatactggc     1800
attgatagga cattaaattg tatgattgga cagatgaagc atattttggc tgcagaacag     1860
aagaaaacag attttaagcc agaagatgaa aacaatgttt tgattcaata tactaatgcc     1920
tgtgtaaaag tctgtgctta cgtaagaaaa caagtggaga agattaaaaa ttccatggat     1980
gggaagaatg tggatacagt tttgatggaa cttggagtac gttttcatcg acttatctat     2040
gagcatcttc aacaatatte ctacagttgt atgggtggca tgttgccat ttgtgatgta     2100
gccgaatata ggaagtgtgc caaagacttc aagattccaa tggattaca tctttttgat     2160
actctgcatg ctctttgcaa tcttctggta gttgcccag ataatttaaa gcaagtctgc     2220
tcaggagaac aacttgctaa tctggacaag aatatacttc actccttcgt acaacttcgt     2280
gctgattata gatctgccg ccttgctcga cacttcagct gagattgaat ttacaaagga     2340
att                                                    2343

```

<210> 369
 <211> 708
 <212> PRT
 <213> Homo sapiens

<400> 369

```

Met Ala Thr Thr Ala Glu Leu Phe Glu Glu Pro Phe Val Ala Asp Glu
 1          5          10          15
Tyr Ile Glu Arg Leu Val Trp Arg Thr Pro Gly Gly Gly Ser Arg Gly
          20          25          30
Gly Pro Glu Ala Phe Asp Pro Lys Arg Leu Leu Glu Glu Phe Val Asn
          35          40          45
His Ile Gln Glu Leu Gln Ile Met Asp Glu Arg Ile Gln Arg Lys Val
          50          55          60
Glu Lys Leu Glu Gln Gln Cys Gln Lys Glu Ala Lys Glu Phe Ala Lys
65          70          75          80
Lys Val Gln Glu Leu Gln Lys Ser Asn Gln Val Ala Phe Gln His Phe
          85          90          95
Gln Glu Leu Asp Glu His Ile Ser Tyr Val Ala Thr Lys Val Cys His
          100          105          110
Leu Gly Asp Gln Leu Glu Gly Val Asn Thr Pro Arg Gln Arg Ala Val
          115          120          125
Glu Ala Gln Lys Leu Met Lys Tyr Phe Asn Glu Phe Leu Asp Gly Glu
          130          135          140
Leu Lys Ser Asp Val Phe Thr Asn Ser Glu Lys Ile Lys Glu Ala Ala
145          150          155          160
Asp Ile Ile Gln Lys Leu His Leu Ile Ala Gln Glu Leu Pro Phe Asp
          165          170          175
Arg Phe Ser Glu Val Lys Ser Lys Ile Ala Ser Lys Tyr His Asp Leu
          180          185          190
Glu Cys Gln Leu Ile Gln Glu Phe Thr Ser Ala Gln Arg Arg Gly Glu
          195          200          205
Ile Ser Arg Met Arg Glu Val Ala Ala Val Leu Leu His Phe Lys Gly
          210          215          220
Tyr Ser His Cys Val Asp Val Tyr Ile Lys Gln Cys Gln Glu Gly Ala
225          230          235          240
Tyr Leu Arg Asn Asp Ile Phe Glu Asp Ala Gly Ile Leu Cys Gln Arg
          245          250          255
Val Asn Lys Gln Val Gly Asp Ile Phe Ser Asn Pro Glu Thr Val Leu
          260          265          270
Ala Lys Leu Ile Gln Asn Val Phe Glu Ile Lys Leu Gln Ser Phe Val
          275          280          285
Lys Glu Gln Leu Glu Glu Cys Arg Lys Ser Asp Ala Glu Gln Tyr Leu
          290          295          300
Lys Asn Leu Tyr Asp Leu Tyr Thr Arg Thr Thr Asn Leu Ser Ser Lys
305          310          315          320
Leu Met Glu Phe Asn Leu Gly Thr Asp Lys Gln Thr Phe Leu Ser Lys
          325          330          335
Leu Ile Lys Ser Ile Phe Ile Ser Tyr Leu Glu Asn Tyr Ile Glu Val
          340          345          350
Glu Thr Gly Tyr Leu Lys Ser Arg Ser Ala Met Ile Leu Gln Arg Tyr
          355          360          365
Tyr Asp Ser Lys Asn His Gln Lys Arg Ser Ile Gly Thr Gly Gly Ile
          370          375          380
Gln Asp Leu Lys Glu Arg Ile Arg Gln Arg Thr Asn Leu Pro Leu Gly
385          390          395          400
Pro Ser Ile Asp Thr His Gly Glu Thr Phe Leu Ser Gln Glu Val Val
          405          410          415

```

Val Asn Leu Leu Gln Glu Thr Lys Gln Ala Phe Glu Arg Cys His Arg
 420 425 430
 Leu Ser Asp Pro Ser Asp Leu Pro Arg Asn Ala Phe Arg Ile Phe Thr
 435 440 445
 Ile Leu Val Glu Phe Leu Cys Ile Glu His Ile Asp Tyr Ala Leu Glu
 450 455 460
 Thr Gly Leu Ala Gly Ile Pro Ser Ser Asp Ser Arg Asn Ala Asn Leu
 465 470 475 480
 Tyr Phe Leu Asp Val Val Gln Gln Ala Asn Thr Ile Phe His Leu Phe
 485 490 495
 Asp Lys Gln Phe Asn Asp His Leu Met Pro Leu Ile Ser Ser Ser Pro
 500 505 510
 Lys Leu Ser Glu Cys Leu Gln Lys Lys Lys Glu Ile Ile Glu Gln Met
 515 520 525
 Glu Met Lys Leu Asp Thr Gly Ile Asp Arg Thr Leu Asn Cys Met Ile
 530 535 540
 Gly Gln Met Lys His Ile Leu Ala Ala Glu Gln Lys Lys Thr Asp Phe
 545 550 555 560
 Lys Pro Glu Asp Glu Asn Asn Val Leu Ile Gln Tyr Thr Asn Ala Cys
 565 570 575
 Val Lys Val Cys Ala Tyr Val Arg Lys Gln Val Glu Lys Ile Lys Asn
 580 585 590
 Ser Met Asp Gly Lys Asn Val Asp Thr Val Leu Met Glu Leu Gly Val
 595 600 605
 Arg Phe His Arg Leu Ile Tyr Glu His Leu Gln Gln Tyr Ser Tyr Ser
 610 615 620
 Cys Met Gly Gly Met Leu Ala Ile Cys Asp Val Ala Glu Tyr Arg Lys
 625 630 635 640
 Cys Ala Lys Asp Phe Lys Ile Pro Met Val Leu His Leu Phe Asp Thr
 645 650 655
 Leu His Ala Leu Cys Asn Leu Leu Val Val Ala Pro Asp Asn Leu Lys
 660 665 670
 Gln Val Cys Ser Gly Glu Gln Leu Ala Asn Leu Asp Lys Asn Ile Leu
 675 680 685
 His Ser Phe Val Gln Leu Arg Ala Asp Tyr Arg Ser Ala Arg Leu Ala
 690 695 700
 Arg His Phe Ser
 705

<210> 370
 <211> 60
 <212> DNA
 <213> Homo sapiens

<400> 370
 gtcaatcact ctcccagcat aagcacccca gccactcta ttccaggag tcatgctatg 60

<210> 371
 <211> 60
 <212> DNA
 <213> Homo sapiens

<400> 371
agtagaattt cctctggaac tggagacatt ttccagcaac atattcagct tgaaagtaca 60

<210> 372
<211> 60
<212> DNA
<213> Homo sapiens

<400> 372
ccagagactg gagatcctgt tacgctgaga ctccctgatg atggagcagg tgctgatgtt 60

<210> 373
<211> 60
<212> DNA
<213> Homo sapiens

<400> 373
ttacagtctg ctgtatctaa cattgccag ggcctctgt ttattccccc caattctgat 60

<210> 374
<211> 60
<212> DNA
<213> Homo sapiens

<400> 374
gctgtgcccc cagccactgt ggaagccttt gtggaaagag acagcctcca ttttctcat 60

<210> 375
<211> 60
<212> DNA
<213> Homo sapiens

<400> 375
aaaaacacag tgactgtgga taatactgtg ggcaacgaca ctatgtttct agttacgtgg 60

<210> 376
<211> 20
<212> PRT
<213> Homo sapiens

<400> 376
Leu Gln Ser Ala Val Ser Asn Ile Ala Gln Ala Pro Leu Phe Ile Pro
5 10 15

Pro Asn Ser Asp
20

<210> 377
<211> 20
<212> PRT
<213> Homo sapiens

<400> 377

Val Asn His Ser Pro Ser Ile Ser Thr Pro Ala His Ser Ile Pro Gly
 5 10 15

Ser His Ala Met
 20

<210> 378

<211> 20

<212> PRT

<213> Homo sapiens

<400> 378

Pro Glu Thr Gly Asp Pro Val Thr Leu Arg Leu Leu Asp Asp Gly Ala
 5 10 15

Gly Ala Asp Val
 20

<210> 379

<211> 20

<212> PRT

<213> Homo sapiens

<400> 379

Ala Val Pro Pro Ala Thr Val Glu Ala Phe Val Glu Arg Asp Ser Leu
 5 10 15

His Phe Pro His
 20

<210> 380

<211> 20

<212> PRT

<213> Homo sapiens

<400> 380

Ser Arg Ile Ser Ser Gly Thr Gly Asp Ile Phe Gln Gln His Ile Gln
 5 10 15

Leu Glu Ser Thr
 20

<210> 381

<211> 20

<212> PRT

<213> Homo sapiens

<400> 381

Lys Asn Thr Val Thr Val Asp Asn Thr Val Gly Asn Asp Thr Met Phe

